

**Regional variation in the realization  
of intonation contours  
in the Netherlands**

Published by  
LOT  
Trans 10  
3512 JK Utrecht  
The Netherlands

phone: +31 30 253 6111

e-mail: [lot@let.uu.nl](mailto:lot@let.uu.nl)  
<http://www.lotschool.nl>

Cover illustration: Judith Hanssen, *Imprints of the waves on a Zeelandic beach.*

ISBN: 978-94-6093-229-8  
NUR 616

Copyright © 2017: Judith Hanssen. All rights reserved.

# **Regional variation in the realization of intonation contours in the Netherlands**

Een wetenschappelijke proeve op het gebied van de Letteren

## **Proefschrift**

ter verkrijging van de graad van doctor  
aan de Radboud Universiteit Nijmegen  
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,  
volgens besluit van het College van Decanen  
in het openbaar te verdedigen op vrijdag 10 maart 2017  
om 14:30 uur precies

door

**Judith Elisabeth Gerarda Hanssen**

geboren op 17 december 1979

te Born

*Promotoren:*

Prof. dr. C.H.M. Gussenhoven (Radboud Universiteit Nijmegen)

Prof. dr. J. Peters (Carl von Ossietzky Universität, Oldenburg)

*Manuscriptcommissie:*

Prof. dr. R. van Hout (voorzitter; Radboud Universiteit Nijmegen)

Prof. dr. M. Grice (Universität zu Köln, Duitsland)

Prof. dr. M. van Oostendorp (Meertens Instituut, Universiteit Leiden)

The research reported in this dissertation was supported by the Netherlands Organization for Scientific Research (NWO), Grant No. 360-7-180 awarded to Prof. dr. C.H.M. Gussenhoven.

## ACKNOWLEDGEMENTS

---

These words of gratitude are the last ones to be written and yet they appear on one of the first pages. Of course that is because without the help of everyone mentioned here, all the other pages would never have existed.

Let me start by expressing my sincere gratitude to my supervisors Carlos Gussenhoven and Jörg Peters for their most valuable advice and for their elaborate comments on drafts and papers. Carlos, this has been a rocky road to say the least. I kept taking unnecessary by-paths, sometimes getting almost out of sight. Nevertheless, you've managed to guide me towards the main road more than once and eventually to the finish line. I greatly appreciate your ongoing support and your incredibly quick responses to my questions no matter the time of day (or night). Jörg, you were my daily supervisor at the initial stages of the project and taught me about statistics and the phonetics of intonation. I enjoyed creating and piloting the experimental materials together, and traveling to various places for the first set of recordings. My work has benefited from your eye for detail time and again. Thank you for staying involved.

At various stages during the project, I have benefited from interesting discussions with, and advice from, various people. Anneke Neijt, Bettina Braun, Frank Kügler, Toni Rietveld and Vincent van Heuven, thank you for sharing your knowledge. Thank you also to the members of the manuscript committee, Roeland van Hout, Martine Grice and Marc van Oostendorp, for taking the time to read my manuscript, and for their helpful comments.

Joop Kerkhoff has provided technical support from the start. Frank van den Beuken, Lian van Hoof, Jan Michalsky, Juul Thijssen and Renske Teeuw have assisted in recording, processing or annotating the speech materials. Thank you all for your contribution to the project. Very special thanks to Rachel Fournier for technical and practical help, but even more so for valuable personal support which has often boosted my confidence.

Many people have contributed to the data collection in Goes, Rotterdam, Amsterdam, Grou and Winschoten, either by offering advice, by translating the materials, by recruiting participants or by providing time and space for the recordings. I would like to thank Pieter Duijff and Wytse Rypma (Fryske Akademy), everyone at Bureau Hooghiemster, Omroep Friesland, Kor Hylkema (De Grouster), Dr. Marron Fort, Hennie Lemein, Marjo van Koppen, Tom Geenen (Scouting Maasgroep 18), Piet Karman, Jan Karman, Marcel van den Borgt (Ostrea Lyceum), and Ingrid and Iwan van Klingereren (Scouting Phoenix).

Of course I am also indebted to all the speakers, for their time and willingness to participate.

Over the years, I have had the pleasure to work with and meet many people. In particular, I want to mention Marjel van Dijk, Lotte Hogeweg, Xuliang He, Yiya Chen, Maïke Prehn, Marco van de Ven and Monique Lamers. Thank you for many enjoyable conversations, (career) advice, and your general interest.

I am grateful to my colleagues at Avans. I feel lucky to be part of such a pleasant, inspiring and supportive team. Particularly, I wish to acknowledge Natasja Nova and Ron Tenge, who have facilitated writing this thesis by providing time and understanding.

A special word of thanks to my parents Truus and Theo, my ‘other’ parents Anny and Jan, my sisters Lilian and Inge, and my brothers and sister in-law for expressing their belief in me, for never ceasing to support me in so many ways, as well as for many hours of babysitting. Thanks also to Annelies, Claudine, Dieuwertje, Eveline, Joanne, Lotte, Marina, Marloes and Renske for knowing when (not) to ask about progress, and for essential moments of fun during camping trips, coffee breaks and dinners. Thanks to Joep and Maarten for agreeing to be my *paranymphs* on top of that. A big thanks to my children Felix, Oskar and Tygo for being patient when I had to work on ‘the book’, and for bringing so much laughter into our home. Finally, I am incredibly grateful to Bas. Thank you for putting up with my nocturnal writing habits for so many years, thank you for your unlimited patience, care (and soup). Without it, and without you, I would not have been able to complete this thesis.

CONTENTS

---

**INTRODUCTION..... 1**

1.1 Intonation in Dutch dialectal research ..... 1

1.2 International studies on dialect intonation ..... 4

1.3 Outline and scope of the thesis ..... 6

References ..... 10

**THEORETICAL FRAMEWORK..... 13**

2.1 The structure of intonation ..... 13

2.2 The realization of phonological structures ..... 17

2.2.1 Tonal context..... 19

2.2.2 Segmental effects on timing..... 21

2.2.3 Prosodic context ..... 22

2.2.4 Pitch variation for paralinguistic purposes ..... 24

2.3 Cross-linguistic variation in intonation ..... 25

2.4 Variation: phonology or phonetics?..... 27

References ..... 28

**REGIONAL VARIATION IN PHONETIC RESPONSES TO TIME PRESSURE IN  
DUTCH IP-FINAL NUCLEAR CONTOURS..... 31**

Abstract ..... 31

3.1 Introduction..... 31

3.1.1 Measuring alignment ..... 37

3.1.2 Phonetics or phonology?..... 38

3.2 Materials and method..... 39

3.2.1 Varieties and subjects ..... 39

3.2.2 Materials ..... 41

3.2.3 Procedure ..... 42

3.2.4 Data selection and analysis..... 43

3.2.5 Statistical and visual analysis ..... 47

3.3 Results..... 47

3.3.1 Sonorant rime duration..... 47

3.3.2 Avoidance of complex contours ..... 49

3.3.3 Durational differences between complex vs simple contours..... 50

3.3.4 Timing adjustments under time pressure..... 50

3.3.4.1 Simple vs complex contours ..... 51

3.3.4.2 Falls ..... 51

3.3.4.3 Fall-rises..... 53

3.3.4.4 Rises ..... 54

3.3.5 Shape adjustments under time pressure..... 55

3.3.5.1 Falls ..... 59

3.3.5.2 Fall-rises..... 62

3.3.5.3 Rises .....	69
3.4 Discussion .....	71
3.4.1 The realization of IP-final complex contours .....	71
3.4.2 Peak / Valley retraction .....	72
3.4.3 Truncation / Compression / Undershoot .....	74
3.5 Conclusion .....	75
References .....	78
<b>PHONETIC EFFECTS OF FOCUS IN FIVE VARIETIES OF DUTCH .....</b>	<b>81</b>
Abstract .....	81
4.1 Introduction .....	81
4.2 Method .....	83
4.2.1 Varieties and subjects .....	83
4.2.2 Materials and procedure .....	84
4.2.3 Data analysis .....	85
4.3 Results .....	87
4.3.1 Segmental duration .....	87
4.3.2 Scaling of tonal targets .....	88
4.3.3 Nuclear contour shape .....	88
4.3.4 Tonal timing .....	90
4.4 Discussion and conclusions .....	90
4.4.1 Hyperarticulation .....	91
4.4.2 Focus size vs type .....	92
4.4.3 Contextual clues .....	92
References .....	93
<b>FINAL AND NON-FINAL NUCLEAR CONTOURS ACROSS VARIETIES OF DUTCH, FRISIAN AND LOW SAXON .....</b>	<b>95</b>
Abstract .....	95
5.1 Introduction .....	95
5.2 Procedure .....	98
5.2.1 Materials .....	98
5.2.2 Varieties and subjects .....	99
5.2.3 Recording procedure and data selection .....	101
5.2.4 Variables and analysis .....	102
5.3 Results .....	104
5.3.1 Sonorant rime duration .....	104
5.3.2 Peak timing .....	106
5.3.3 Scaling of tonal targets .....	108
5.3.4 Contour shape: f0 duration, excursion and slope .....	110
5.3.4.1 Non-final and final falls .....	110
5.3.4.2 Fall-rises .....	112
5.4 Summary and discussion .....	114
5.4.1 Effects of gender .....	115
5.4.2 Effects of sentence condition .....	115



5.4.3	Effects of dialect .....	116
	References .....	121
<b>NON-STANDARD MELODIES AND MELODY PREFERENCES IN DIALECT-ACCENTED DUTCH.....</b>		<b>125</b>
	Abstract .....	125
6.1	Introduction.....	125
6.2	Procedure .....	127
6.2.1	Materials .....	127
6.2.2	Participants .....	128
6.2.3	Recordings .....	130
6.3	Nuclear tone preferences in different pragmatic contexts .....	130
6.3.1	Labels .....	130
6.3.2	Results.....	131
6.3.2.1	Statements .....	131
6.3.2.2	Yes/no-questions.....	132
6.3.2.3	Rhetorical questions .....	137
6.3.3	Discussion.....	137
6.4	Dialect-specific contour realizations.....	139
6.4.1	Falls in Rotterdam and Amsterdam.....	139
6.4.2	Zuid-Beveland declarative and interrogative falls.....	147
6.4.3	Zuid-Beveland ‘rise-rise’ .....	149
6.4.4	Discussion.....	151
6.5	Conclusion .....	153
	References .....	154
<b>SUMMARY AND CONCLUSIONS .....</b>		<b>157</b>
<b>APPENDIX .....</b>		<b>167</b>
<b>SAMENVATTING IN HET NEDERLANDS.....</b>		<b>179</b>
<b>CURRICULUM VITAE .....</b>		<b>183</b>



### 1.1 Intonation in Dutch dialectal research

When I tell people that my dissertation is about the melody of Dutch dialects, they often respond by saying how they find the intonation of dialect X ‘funny’, or how dialect Y ‘sings’. It is a popular belief that dialects have their own characteristic melodies, and that this intonation is an important cue to a speaker’s (linguistic and geographical) origins. Indeed, feeding the words *dialect* and *zangerig* (‘lilting’ or ‘sing-songy’) to search engine Google suggests that quite a number of dialects in the Netherlands are considered to sound more melodious than others. The Limburgian dialects are most often mentioned as sounding melodious, but the dialects of Zeeland, Rotterdam, Leiden, Scheveningen and the Zaanstreek are also associated with sing-song intonation.

Although many popular beliefs on language are subjective (such as comments that some dialects sound *nicer* than others), in this case people may be right in implying that the melody of speech varies from one dialect to another. Some evidence that this is true is provided by Gooskens (1997), who explored whether or not prosodic information (i.e., pitch, duration and loudness, Gooskens 1997:1) plays a role in the perception of dialectal differences. The results of the perception experiments suggested that both exogenous listeners (Standard Dutch speakers) and endogenous listeners (dialect speakers) used prosodic information to distinguish their own variety from other dialects. It is not the case that speakers of Standard Dutch were able to identify language varieties in the absence of verbal information (i.e., syntax, lexicon, morphology, and segmental phonetics / phonology, Gooskens 1997:1), but prosodic information improved language identification if it was combined with verbal information. A stronger case for the relevance of prosodic information was provided by the identification scores of native speakers of the investigated dialects. When dialect speakers were presented with non-verbal speech materials of Standard Dutch and their own variety, they could identify their own variety above chance level.

Despite the widespread idea that varieties of Dutch differ with respect to intonation, and despite Gooskens’ results which suggest that this is indeed the case, the study of regional intonation in the Netherlands has been restricted to the dialects of the South-East (the province of Limburg). Most Limburgian dialects use tone at the word level (lexical tone) and the sentence level (intonation). A recent large-scale project at Radboud University Nijmegen

investigated the production, perception, and historical development of lexical tones, and their interaction with intonational melodies, for several Limburgian varieties spoken in the Netherlands and Belgium (see, e.g., Gussenhoven and van der Vliet 1999, Gussenhoven 2000, Hanssen 2005, Fournier et al. 2006, Peters 2007, Fournier 2008, Gussenhoven and Peters 2008, Gussenhoven 2012 and Gussenhoven and van den Beuken 2012). However, we still know next to nothing about the intonation of dialects outside Limburg. It is not the case that linguists are simply not interested in Dutch dialects. On the contrary, the study of linguistic and cultural variation in the Netherlands flourishes in KNAW institutes like the Meertens Institute in Amsterdam and the Fryske Akademy in Leeuwarden. The Meertens Institute has documented dialectal variation in three atlases covering morphological, syntactic and phonological variation in over 500 places (MAND, SAND and FAND). The Fryske Akademy has published a comprehensive dictionary of the Frisian language, while the NCDN (Nijmegen Centre for Dialectology and Onomastics) has published multi-volume dictionaries of the dialects of Brabant, Gelderland and Limburg. The thriving study of Dutch dialects is also evident from the popular multi-volume series *Taal in stad en land* ('Language in city and country', van der Sijs 2002-2005) that describes 27 regional varieties of Dutch and Frisian, regional speech databases such as *De Nederlandse Dialectenbank* (the Dutch dialect database)<sup>1</sup>, the Corpus Spoken Frisian<sup>2</sup>, and numerous dialect-related websites<sup>3</sup>.

Quite in contrast with the popularity of studying dialectal variation, references to intonation are hard to come by. This void should not be attributed to a lack of interest in intonation on the part of Dutch dialectologists. In fact, more than sixty years before Gooskens, van Es (1932:93) already suggested that intonation may be the most important cue to someone's origins, and that a dialect may have specific intonation characteristics. He later illustrates this with an impressionistic comparison of the intonation of two question types in Katwijk and Frisian dialect (van Es 1935). Van Es stresses the importance of studying the intonation of dialects because they can contribute to our understanding of the development of the standard language. Another early reference to prosodic variation in Dutch dialects is Daan (1938:473), who similarly says that

“[...] it is possible to recognize a dialect by other characteristics than only by the words and sounds, [but] it is very difficult to say which these characteristics are. No doubt the musical accent plays an important part in this matter”.

Based on both spontaneous and read speech, Daan provides impressionistic prosodic characteristics of some regional varieties of Dutch. She claims that

---

<sup>1</sup> [www.meertens.knaw.nl/ndb/](http://www.meertens.knaw.nl/ndb/).

<sup>2</sup> [www1.fa.knaw.nl/ksf.html](http://www1.fa.knaw.nl/ksf.html).

<sup>3</sup> visit *streektaal.net* at <http://taal.phileon.nl/links.php> as a starting point.

speakers of Zeeland, Noord-Holland and Friesland have ‘excessive musical accent’, whereas those of West-Brabant, Zuid-Holland, Utrecht and the Northeast of the Netherlands have ‘normal pitch patterns’. She calls the dialect of Drenthe monotonous, says that slowness is probably typical of the North Holland dialects, and claims that Frisian is characterized by big intervals and rising at the end [of an utterance, JH].

Almost thirty years later, Weijnen (1966:279) devotes only half a page to dialect intonation in his comprehensive handbook on Dutch dialectology. His comments are restricted to a few informal and impressionistic characterizations. Some dialects (e.g., Limburgian, Northern Hollandic, and the dialect of Dirksland in Zuid-Holland) are described as having lilting intonation, whereas others, like the Groningen dialect, are explicitly said not to be melodious. Finally, Frisian is characterized as having a tendency to display very idiosyncratic melodic sounds.

More recently, both Wortel (2002) and van Oostendorp (2002) make reference to intonation in their descriptions of the Leiden and Rotterdam dialect, respectively, which appeared in the series *Taal in stad en land*. Wortel (2002:86,89) writes two passages on tonal raising in the Leiden dialect, which he connects to the pronunciation of /r/. Van Oostendorp (2002) mentions that the Rotterdam dialect has a characteristic intonation which is often referred to as ‘lilting’. He suggests that the intonation of Rotterdam and many other dialects is characterized as lilting because we compare dialect intonation to the melody of Standard Dutch, which he describes as relatively flat. Van Oostendorp (2002:37) regrets that his comments are necessarily based on his own impressions, because

“unfortunately no scientific research has been conducted on the differences in intonation of the different city dialects”.

This quote by van Oostendorp applies not only to Dutch city dialects, but to all non-tonal regional varieties of the Netherlands. The descriptions of Dutch dialect intonation summarized above are not based on detailed acoustic analyses and are hence impressionistic in nature. Reasons for the lack of attention to intonation in Dutch dialectological research can be traced back to van Es’ (1932) point that the characterization of dialects in terms of intonation requires *recording and analyzing many sentence types by many speakers*. Analysis of intonation should be based on speech materials, which is why the use of questionnaires, which can be more easily processed in large numbers, is not a suitable tool for this type of research. Detailed analysis of intonation is a complex and time-consuming task, and indeed Gooskens (1997:3) acknowledges that until the relatively recent development of computerized methods of analysis, it was difficult to describe prosodic variation in a consistent way. In the decade after Gooskens’ dissertation appeared, we have seen great advances in technical applications that facilitate the recording, processing, annotation and analysis of large data sets of speech materials. Moreover, from the early 80’s, the insight has gained ground that intonation is structured, which resulted in the development of a theoretical

framework for the description of intonation that can be applied to all tonal and non-tonal languages (the Autosegmental-Metrical (AM) framework, see Chapter 2 of this dissertation for a concise introduction, or Ladd 2008:43-84 and Gussenhoven 2004:123-142 for a detailed description). Since that time, studies on intonation constitute an increasing part of phonological and phonetic research. The grammar of Standard Dutch intonation has been described using the transcription model ToDI (Transcription of Dutch Intonation, Gussenhoven et al. 2002, Gussenhoven 2005). This model has been used for all transcriptions in this thesis and will also be introduced in Chapter 2.

## 1.2 International studies on dialect intonation

Cross-linguistic studies have revealed that intonation can vary in a number of ways. To cite Ladd (2008:115), there are ‘different types of differences’. Analogous to Wells’ (1982) classification set for segmental phonetic differences<sup>4</sup>, Ladd (2008:116) proposes that cross-linguistic variation in intonation can be described along four dimensions: variation can be categorized as (1) semantic, (2) systemic, (3) phonotactic and (4) realizational. Languages may share the same set of melodies but use them in different situations or to express different meanings (*semantic variation*). It is also possible that a particular melody occurs in one language, but not in another (*systemic variation*). A third category is *phonotactic variation*, which is related to differences in the way languages combine high and low tones to form intonational melodies. Finally, a shared melody may be pronounced slightly differently in two languages, or may be realized differently depending on the context in which it occurs (*realizational variation*). These four categories are described in more detail in Chapter 2.3.

Note that each instance of the word ‘language’ in the previous paragraph can be replaced by the word ‘variety’ or ‘dialect’. Over the past decades, there has been an increasing interest not only in intonational phonology, but also in cross-dialectal variation in intonation. This interest is reflected in an increasing number of large-scale projects, along with some smaller-scale investigations and mono-dialectal studies, that are concerned with this topic. A list of projects, restricted to European languages, includes:

---

<sup>4</sup> Which in turn is based on Trubetzkoy (1931), as Peters (2006:70) points out.

- *English Intonation in the British Isles*  
Grabe, Nolan, Farrar, Post, 1997-2002<sup>5</sup>
- *SweDia 2000, The phonetics and phonology of the Swedish dialect around the year 2000*  
Bruce, Elert, Engstrand, Bannert, 1998-2004<sup>6</sup>
- *Studies on the Structure and Function of Regional Intonation Contours in German*  
Auer, Selting, Gilles, Koeser, Peters, 1998-2004<sup>7</sup>
- *Tonal varieties of Dutch: Structure, Function, and Perception*  
Gussenhoven, Peters, Fournier 2003-2006<sup>8</sup>
- *Prosody of Irish Dialects: the use of intonation, rhythm, voice quality for linguistic and paralinguistic signalling*  
Ní Chasaide, Dalton, 2003-2006<sup>9</sup>
- *The Interactive atlas of Catalan/Occitan/Spanish intonation*  
Prieto and colleagues, 2008-2011<sup>10</sup>

The results of all these projects have contributed to the current understanding that intonation shows differences even between varieties of the same language.

All in all, the time is right to start describing and documenting the intonational patterns of regional varieties in the Netherlands. The two main questions to be answered are:

1. whether Dutch dialects have their own characteristic intonation, or ‘signature’, and
2. if so, what this intonation sounds like and how it differs from intonation in other varieties.

The research programme *Intonation in Varieties of Dutch*<sup>11</sup> aims to provide an answer to these questions by recording varieties in the Netherlands and North-

---

<sup>5</sup> [http://www.phon.ox.ac.uk/files/apps/old\\_IViE/](http://www.phon.ox.ac.uk/files/apps/old_IViE/), and e.g. Grabe 2004.

<sup>6</sup> <http://swedia.ling.gu.se>.

<sup>7</sup> [http://paul.igl.uni-freiburg.de/auer/?Forschung:Abgeschlossene\\_Drittmittelprojekte:Dialektintonation](http://paul.igl.uni-freiburg.de/auer/?Forschung:Abgeschlossene_Drittmittelprojekte:Dialektintonation) and e.g. Peters et al. 2015.

<sup>8</sup> <http://www.lotpublications.nl/perception-of-the-tone-contrast-in-east-limburgian-dialects-perception-of-the-tone-contrast-in-east-limburgian-dialects>, <http://www.clips.ua.ac.be/projects/tonal-dialects-in-dutch-structure-perception-and-function>, and e.g. Gussenhoven 2000.

<sup>9</sup> <http://www.tcd.ie/slscs/research/projects/past/prosody.php>, and e.g. Dalton and Ní Chasaide 2006.

<sup>10</sup> <http://prosodia.upf.edu/atlesentonacio/index-english.html>, and e.g. Prieto, Estebas, and Vanrell 2010.

<sup>11</sup> NWO Grant no. 360-70-180 awarded to prof. dr. C.H.M. Gussenhoven (2006-2010).

West Germany, including Standard Dutch, non-native Dutch, and six local varieties (including Frisian). The local varieties are spoken along the entire Dutch North Sea coast, extending to North-West Germany (see Figure 1). Sub-programmes include investigations of prosodic aspects of Mandarin-accented Dutch (He, 2012), which aims to contribute to the understanding of prosodic transfer from a typologically very different language like Chinese to Dutch. A second sub-programme is concerned with intonational melodies that are typically found in one variety but not the other (e.g., Peters and Gussenhoven, ms).

This dissertation reports the results of a sub-programme describing *phonetic variation between intonation contours in varieties of Dutch*, by recording and analyzing carefully constructed speech materials from six varieties spoken in the Netherlands only: *Zeelandic* as spoken in Zuid-Beveland, the two urban *Hollandic* varieties in Rotterdam and Amsterdam, *Frisian* in Grou, *Low Saxon* in Winschoten, and finally *Standard Dutch*. In addition to presenting phonetic (i.e., non-structural) variation, this thesis also touches upon phonological differences between the varieties studied as they emerged from the experiments that were conducted.



Figure 1. Recording locations in the Netherlands, including Weener in North-West Germany, which is not analyzed in this dissertation.

### 1.3 Outline and scope of the thesis

From the outset, the main goal of this study was to investigate whether and how varieties of Dutch show phonetic variation in the realization of (phonologically) identical sentence melodies. Earlier studies have reported that the realization of intonation contours may depend on the context in which they occur, and that, moreover, languages may respond differently to such contextual variation. Factors that may affect pronunciation are, e.g., the segmental make-up of the



word on which a falling or rising melody is pronounced, the number of syllables over which a melody is pronounced, or the sentence position of the word that carries the intonation contour. Work on the latter topic by Grabe (1998) has shown that, when a sentence accent is pronounced near the end of the utterance, some speakers realized the melody at a higher speed (*compression*) whereas others simply ended the melody prematurely (*truncation*). She showed that these responses to time pressure depended on the intonation contour involved (falling or rising) and were also language-dependent (English vs. German), and even variety-dependent (varieties of British English, Grabe et al. 2000). Chapter 3 investigates how six varieties of Dutch respond to sentence-final time pressure on the realization of three intonation contours: falls, rises and fall-rises. The chapter also reviews and reports other possible responses to time pressure besides truncation and compression, such as the phonological response of avoiding complex contours utterance-finally. Thirdly, it reports whether responses to time pressure are contour- and/or variety-dependent.

Next, Chapter 4 is similarly concerned with phonetic variation. However, we now look at the realization of sentence melodies in relation to the information structure of utterances. In West Germanic, intonation is closely linked with the pragmatic purpose of marking information structure, whereby new information is often marked with a pitch accent, and given information may be deaccented. In the following examples, the focus constituent is marked by square brackets, and the location of the pitch accent is marked by capitals.

- (1) *What's Felix doing?*  
He's [playing with his LEGO train]<sub>FOC</sub>.
- (2) *What is Felix playing with?*  
He's playing with [his LEGO train]<sub>FOC</sub>.
- (3) *Is Felix playing with his wooden train?*  
He's playing with his [LEGO]<sub>FOC</sub> train.

The examples illustrate that a single pitch accent can mark focus constituents of different sizes (e.g., *broad* vs *narrow*). Besides their size, focus constituents can also vary in type. Whereas examples (1) and (2) express '*informational focus*', providing requested information, sentence (3) is an example of '*corrective focus*', correcting a piece of given information.

The sentences in (1) - (3) can be realized with phonologically identical intonation contours, in which case they are ambiguous with respect to focus type and size. However, the word LEGO in the last example can be said to carry a higher *information value* than in the second example. Languages like English and German have been reported to express such pragmatic differences in the realization of the pitch accents. There are many indications that the acoustic prominence of pitch accents is raised as the scope of the focus constituent narrows, and when

the focus meaning is contrastive (cf. (3)), as opposed to non-contrastive (cf. (1) - (2)). Acoustic prominence may be increased by, e.g., changes in the duration of the segmental material in and around the focus constituent, or by varying peak height, peak timing and pitch excursion, but the cues that are used vary between languages and language varieties. The experiment reported in Chapter 4 investigates how focus type and the size of the focus constituent affect the realization of falling intonation contours in six varieties of Dutch.

Chapter 5 is concerned with what we may call *language-specific phonetic variation*. Even though the experiments reported in Chapters 3 and 4 were specifically designed to study context-specific realizational variation (as a result of differences in time pressure and information value), the data also allowed us to investigate whether the six varieties show more general differences in the realization of intonation patterns. Language-specific realizational differences have been reported to exist between English and Dutch, for example. British English speakers use a significantly wider pitch range than Dutch speakers (Willems 1982, de Pijper 1983). Such differences have also been reported between varieties of the same language, as Atterer and Ladd (2004) have shown for Northern and Southern German, which differ with respect to the timing of pitch accents that occur in non-final sentence position. In Chapter 5, I report language-specific realizational differences between six varieties of Dutch, based on the data collected for Chapters 3 and 4. Such differences may contribute to the widespread idea, described in 1.1 above, that some varieties in the Netherlands have very salient intonational characteristics. If we find evidence in support of this idea, it will be interesting to see whether the intonational distance between varieties changes along with their geographical distance and gradually shifts from the South-West to the North-East of the Netherlands.

In Chapter 6, I partly step away from phonetics and report phonological (i.e., systemic, semantic and/or phonotactic) differences between the varieties as they emerged from the data. This chapter also describes dialect-specific intonation patterns that were in some way or another ‘deviant’ from other varieties. Once again, I should mention that the data were not designed specifically for this purpose. Not all participants realized the sentences in the reading tasks with the same intonation contours, or with the intonation contour the author had expected them to use. Instead of discarding these data and treating them as ‘noise’, I believe it is worth reporting the variation and critically study it. For, surely, the choices speakers make for a particular intonation contour may contribute to the intonational characteristics of their dialect just like the realization of these contours do. As such, the work reported in this chapter adds answers to the questions raised earlier.

Chapter 7 summarizes the main results of the research and contains the discussion and conclusion. In this final chapter, I will look back at the two main questions raised in the introduction. I will reflect on the work carried out for this

dissertation, as well as present suggestions and recommendations for future research.

Before the results of the reading tasks are presented, Chapter 2 will outline the theoretical framework within which this research should be placed, along with an account and discussion of theoretical issues that underlie the investigations in the chapters that follow. This chapter might be of interest to the reader with a background in dialectology who is perhaps less familiar with the phonology and phonetics of intonation.

Chapters 3 to 6 are all (slightly) adapted versions of either published, accepted or submitted manuscripts, and are as such stand-alone texts that can be read in isolation. Due to the fact that the materials reported in all four chapters were collected on the same occasion and from the same set of participants, there is inevitably some overlap between the manuscripts, particularly where the description of the method is concerned.

## References

- Atterer, M. and Ladd, D.R. (2004). On the phonetics and phonology of “segmental anchoring” of F0: evidence from German. *Journal of Phonetics* 32, 177–197.
- Daan, J. (1938). *Dialect and pitch pattern of the sentence*. Proceedings of the 3rd International Congress of Phonetic Sciences, Ghent, 473-480.
- Dalton, M. and Ní Chasaide, A. (2006). Melodic alignment and micro-dialect variation in Connemara Irish. In: C. Gussenhoven and T. Riad (Eds.), *Tones and tunes. Volume 2: Experimental studies in word and sentence prosody*. Berlin: Mouton de Gruyter, 293-315.
- de Pijsper, J. (1983). *Modelling British English intonation*. Dordrecht: Foris.
- Fournier, R. (2008). *Perception of the tone contrast in East Limburgian dialects*. PhD thesis. Utrecht: LOT Publications.
- Fournier, R., Verhoeven, J., Swerts, M., and Gussenhoven, C. (2006). Perceiving word prosodic contrasts as a function of sentence prosody in two Dutch Limburgian dialects. *Journal of Phonetics* 34, 29-48.
- Gooskens, C. (1997). *On the role of prosodic and verbal information in the perception of Dutch and English language varieties*. Wageningen: Ponsen en Looijen.
- Grabe, E. (1998). Pitch accent realization in English and German. *Journal of Phonetics*, 26, 129-144.
- Grabe, E. (2004). Intonational variation in urban dialects of English spoken in the British Isles. In: Gilles, P. and Peters, J. (Eds.), *Regional variation in intonation*. Tübingen: Niemeyer, 9-31.
- Grabe, E., Post, B., Nolan, F., and Farrar, K. (2000). Pitch accent realization in four varieties of British English. *Journal of Phonetics*, 28, 161-185.
- Gussenhoven, C. (2000). The lexical tone contrast of Roermond Dutch in Optimality Theory. In: Horne, M. (Ed.), *Intonation: Theory and Experiment*. Dordrecht: Kluwer, 129-167.
- Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge: Cambridge University Press.
- Gussenhoven, C. (2005). Transcription of Dutch intonation. In: Jun, S.-A. (Ed.), *Prosodic typology: The phonology of intonation and phrasing*. Oxford: Oxford University Press, 118-145.
- Gussenhoven, C. (2012). Asymmetries in the intonation system of Maastricht Limburgish. *Phonology* 29, 39-79.
- Gussenhoven, C. and van der Vliet, P. (1999). The phonology of tone and intonation in the Dutch dialect of Venlo. *Journal of Linguistics*, 35, 99-135.
- Gussenhoven, C., Rietveld, T., Kerkhoff, J., and Terken, J. (2002). *Transcription of Dutch Intonation* (2nd ed.). Available from: <http://todi.let.kun.nl/ToDI/home.htm>.
- Gussenhoven, C. and Peters, J. (2008). De tonen van het Limburgs. *Tijdschrift voor Nederlandse Taalkunde* 13, 87-114.
- Gussenhoven, C. and van den Beuken, F. (2012). Contrasting the high rise and the low rise intonations in a dialect with the Central Franconian tone. *The Linguistic Review* 29, 75-107.
- Hanssen, J. (2005). *Tone and intonation in the dialect of Sittard*. MA thesis, Radboud University Nijmegen.
- He, X. (2012). *Mandarin-accented Dutch Prosody*. PhD thesis. Utrecht: LOT Publications.

- Hoekstra, J. (1991). Oer ist beklamjen fan ferhaldingswurden yn it frysk, it hollansk en int engelsk. *Us Wurk* 40, 67-103.
- Ladd, D.R. (2008). *Intonational phonology* (2<sup>nd</sup> ed.). Cambridge: Cambridge University Press.
- Peters, J. (2006). *Intonation deutscher Regionalsprachen*. Berlin: Walter de Gruyter.
- Peters, J. (2007). A bitonal lexical pitch accent in the Limburgian dialect of Borgloon. In: Riad, T. and Gussenhoven, C. (Eds.), *Tones and tunes. Volume I: Typological studies in word and sentence prosody*. Berlin: Mouton de Gruyter, 167-198.
- Peters, J., Auer, P., Gilles, P., and Selting, M. (2015). Untersuchungen zur Struktur und Funktion regionalspezifischer Intonationsverläufe im Deutschen. In: Kehrein, R., Lameli, A., and Rabanus, S. (Eds.), *Regionale Variation des Deutschen. Projekte und Perspektiven*. Berlin and Boston: De Gruyter Mouton, 53-80.
- Peters, J. and Gussenhoven, C. (ms). *Intonational variation in the Netherlands and beyond. A study of Zeelandic, Hollandic Dutch, West Frisian, Dutch Low Saxon, German Low Saxon, and High German*.
- Prieto, P., Estebas-Vilaplana, E., and Vanrell, M.M. (2010). The relevance of prosodic structure in tonal articulation. Edge effects at the prosodic word level in Catalan and Spanish. *Journal of Phonetics* 38, 687-705.
- Tiersma, P.M. (1999). *Frisian Reference Grammar* (2<sup>nd</sup> ed.). Leeuwarden: Fryske Akademy.
- Trubetzkoy, N.S. (1931). Phonologie und Sprachgeographie. Travaux du Cercle linguistique de Prague 4, 228-234. Reprinted in: Trubetzkoy, N.S. (1939). *Grundzüge der Phonologie*. Prague. 7th edition 1989, Göttingen: Vandenhoeck and Ruprecht, 262-268.
- van Es, G.A. (1935). Syntactische functies der intonatie in de volkstaal onzer noordelijke provinciën. *Handelingen van het Zestiende Nederlandsche Philologen-Congres*, 39-42.
- van Oostendorp, M. (2002). *Rotterdams*. Den Haag: Sdu Uitgevers.
- Weijnen, A. (1966). *Nederlandse dialectkunde*. Assen: Van Gorcum.
- Wells, J.C. (1982). *Accents of English, vol. I: An introduction; vol. II: The British Isles*. Cambridge: Cambridge University Press.
- Willems, N. (1982). *English intonation from a Dutch point of view: An experimental-phonetic investigation of English intonation produced by Dutch native speakers*. H.-I. Ambacht: Intercontinental Graphics
- Wortel, D. (2002). *Leids*. Den Haag: Sdu Uitgevers.



Section 2.1 of this chapter introduces the theoretical model used in this work, the AUTOSEGMENTAL-METRICAL model, and then introduces the transcription system ToDI (Transcription of Dutch Intonation), which is used in all the reports. In section 2.2, it is explained that the realization of intonational structures is determined by (a) structural factors such as tonal context (2.2.1), segmental composition (2.2.2) and prosodic context (2.2.3), which can be captured by *phonetic implementation rules*, and (b) by numerous non-phonological, *paralinguistic* factors (section 2.2.4). Section 2.3 describes four dimensions along which intonation varies cross-linguistically. Finally, I briefly go into the distinction between phonology and phonetics in section 2.4.

### 2.1 The structure of intonation

Intonation concerns the sentence melody of utterances. In languages such as English and Dutch, pitch variation in the sentence functions to signal pragmatic meanings as well as the division into intonational phrases. Intonation may mark differences in discourse meaning such as statements, questions, or non-finality. It also plays an important role in signaling information structure, by marking the prominence relations of an utterance, such as the start of a new contribution to a conversation. Importantly, pitch variation for intonational use is morphological and phonological: it can be described in terms of a set of meaningful contrastive morphemes encoded in terms of discrete phonological segments, called “tones”. While it is generally accepted that pitch variation at the lexical level (in tone languages) is structural, the understanding that intonation has a phonology is relatively recent and is still not generally shared. This may be due to the fact that pitch variation also has a place in a set of paralinguistic gestures, signaling information about the speaker’s attitude towards the listener or the message, or his or her emotional state<sup>1</sup>. Paralinguistic pitch variation is typically non-structural but shares an acoustic channel (fundamental frequency, or  $f_0$ ) with the structural use of pitch variation for intonation purposes, with which it interacts. This interaction challenges anyone involved with the analysis of intonation, who has to separate the linguistic from the paralinguistic use. Secondly, the

---

<sup>1</sup> Extralinguistically, pitch signals biological/physical information, such as the speaker’s sex or age.

continuously variable nature of the acoustic signals that are involved in intonation (f<sub>0</sub>, but also intensity and duration) can make it difficult to recognize the meaningful contrastive elements (tones).

### **Autosegmental-Metrical model**

In the now mainstream AUTOSEGMENTAL-METRICAL (AM) model of intonation<sup>2</sup>, as introduced by Pierrehumbert (1980), intonation contours are described in terms of two categories, the tones H and L. These tones form a linear string that make up the tune of an utterance. Following Autosegmental Phonology (cf. Leben 1973, Goldsmith 1976), the elements of the intonational structure (the tones) are represented on a separate tier from the vowels and consonants on the segmental tier. I follow the assumption in Gussenhoven (2004:144) that each tone within a tune is *aligned*<sup>3</sup>, either with the left or right edge of a prosodic constituent or with another tone, in order to have a location in the structure, but that not every tone is *associated*, which is the coincidental timing of an element on the tonal tier with an element on the segmental tier, a tone bearing unit (TBU). Together, these assumptions explain that there doesn't need to be a one-to-one relation between tones and TBUs, e.g. syllables. In other words, syllables may be tonally underspecified, and conversely not every tone is associated with a syllable.

In languages such as English and Dutch, the tones of a tune are organized into *pitch accents* and *boundary tones*. Pitch accents are aligned with the element in the prosodic hierarchy that is marked for accent, whereas boundary tones align with the edges of some higher-level prosodic constituent like the intonational phrase (IP). The function of the pitch accent varies between languages. In English and Dutch, it is used to mark accented syllables, and is either monotonal (L\* or H\*) or bitonal (e.g., H\*L or LH\*). The star is used to indicate that the tone to its left in the pitch accent associates with the accented syllable. The preceding and following tones in bitonal pitch accents are 'leading' and 'trailing', respectively. These tones are aligned with the starred tone of the pitch accent, but in English and Dutch do not associate with an element on the segmental string.

Boundary tones obviously play a role in signaling the prosodic phrasing of an utterance by demarcating the beginning and end of prosodic phrases such as the intonational phrase (IP), but also function to signal pragmatic meanings like continuation or interrogativity. They are marked by a '%' and are typically

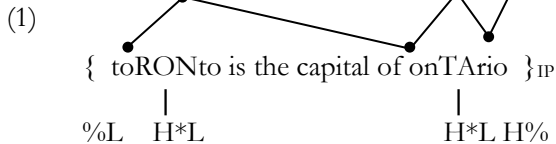
---

<sup>2</sup> Ladd (1996) introduced the term AUTOSEGMENTAL-METRICAL model. For a historical background of AM theory, I refer the reader to Gussenhoven (2004), Ladd (1996, 2008) and Horne (2000), where also information on alternative models can be found.

<sup>3</sup> The term *alignment* (or *phonetic alignment*) is often used to refer to the timing of a tonal target relative to an element in the segmental string, when the realization of contours is described. To avoid confusion, I will use the term *timing* to refer to the phonetic description and the term *alignment* for the structural relationship.



An utterance in English and Dutch is ungrammatical unless there is minimally one pitch accent<sup>4</sup>. A number of prenuclear pitch accents may precede the last (obligatory) nuclear pitch accent. Example (1) (adapted from Gussenhoven 2004:135) presents an English utterance that consists of an intonational tune transcribed with two pitch accents (H\*L) and initial (%L) and final (H%) boundary tones. The realization is presented in idealized form with bullets representing the pitch targets and lines representing the pitch movements in between.



# ToDI - Transcription of Dutch Intonation

<sup>4</sup> Although lower-level prosodic phrases may remain accentless, as in Dutch (Gussenhoven 2005).

between the two analyses lies in the way bitonal pitch accents are analyzed. In a ToBI-style analysis, the contour leading towards the accentual target is analyzed as the pitch accent (*on-ramp* analysis), whereas ToDI follows earlier intonational models (e.g., O'Connor and Arnold 1973) in taking the contour leading off it as constituting the pitch accent (*off-ramp* analysis)<sup>5</sup>. Finally, Gussenhoven proposes a set of phonological rules that modify the underlying representation of pitch accents, which is different from e.g. Pierrehumbert (1980:3), who in her analysis of American English assumes only one phonological level of representation. The differences between the two approaches to analyze the phonology of intonation are discussed in detail in Gussenhoven (2016).

In Gussenhoven's (2005) analysis, the Standard Dutch tonal grammar consists of the following elements:

Initial boundary tones	%H	
	%L	
	%HL	
Final boundary tones	H%	
	L%	
	0%	
Pitch accents	H*	
	L*	
	H*L	
	L*H	
	H*!H	
Modified pitch accents	!H*	Downstep
	!H*L	Downstep
	L*HL	[DELAY]
	L*!HL	Downstep and [DELAY]
	H*LH	<i>prefinal</i> sharp fall

Note that final boundary tones are optional in Standard Dutch (i.e., there is a contrast between L%, H% and 0%), and that a boundary tone complex %HL is assumed in initial position. Pitch accents may be monotonal, bitonal or tritonal. The latter are generated by the phonological rule [DELAY] which operates on the underlying representation of pitch accents and attaches a low tone (L\*) to the left of the pitch accent. In her analysis of American English, Pierrehumbert assumes only one phonological level of representation (Pierrehumbert 1980:3). However,

---

<sup>5</sup> The final pitch accent H\*L would be transcribed L+H\* in ToBI.

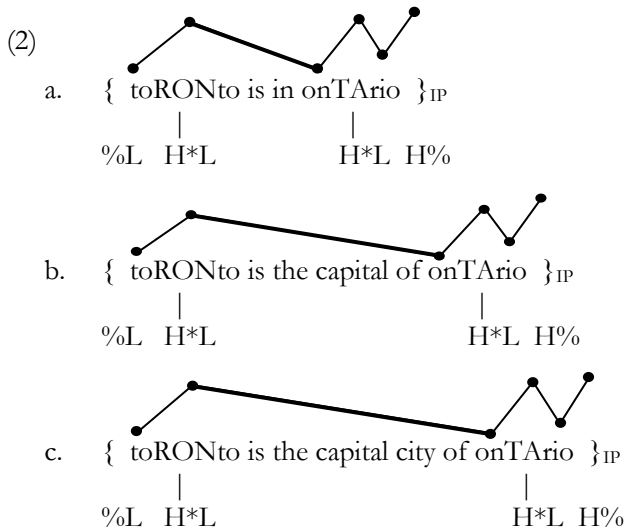
<sup>6</sup> The notation of the absence of a final boundary tone as '0%' is taken from Grabe (1998) and is adopted in this dissertation.

as for Dutch, Gussenhoven (2004) proposes phonological rules that modify the underlying representation of pitch accents in his account of British English. Gussenhoven assumes that all final boundary tone and nuclear pitch accents combinations are well-formed (although not necessarily equally frequent) in Dutch, generating 24 possible nuclear contours. For a full illustration of each of these contours, the reader is referred to Gussenhoven (2005) and He (2012).

With the tonal grammar of a language, intonation structures can be built. This structure is passed onto the next stage, the implementation, and will finally be realized by the speaker.

## 2.2 The realization of phonological structures

Intonation has phonological structure, which provides the input for the phonetic realization. This is why intonation structures that are intuitively the same may sound quite different. The tones that form the phonological structure are mapped to the acoustic signal through a set of phonetic implementation rules, which first determine the timing and scaling of tonal targets. The  $f_0$  movements (the falling and rising contours) between those targets are not phonologically specified but are determined by *linear interpolation*. This rule is important in explaining how the shape of movements can vary without varying the phonological structure. Examples (2a-c) illustrate that the different shapes of the falling pre-nuclear pitch accents are readily explained if we assume that the  $f_0$  movement is determined by interpolation from the pre-nuclear  $H^*$  to the (right-aligning)  $L$ .



Of course, if the  $f_0$  movements are determined by interpolation, we need to specify the realization of the targets at the beginning and end of each movement. The realization of tonal targets can be captured in terms of their location in the segmental structure (*timing*) and their  $f_0$  value (*scaling*).

The timing and scaling are determined by the phonetic implementation, which takes the phonological representation as its input. The scaling of a H-tone will depend on, e.g., its position in the IP (later or earlier), whether it is downstepped, and how much time is available for its realization. The timing of targets is similarly determined by their location in the phonological structure. A pitch accent that is associated with an accented syllable is likely to occur in the vicinity of that syllable. In addition, the use of the asterisk in bitonal accents suggests that the starred tone occurs on the accented syllable. The leading tone will precede the starred tone, and the trailing tone will follow it at some distance<sup>7</sup>. How these tones are timed relative to the starred tone may vary across dialects and languages. Finally, boundary tones that are aligned with IP-edges will be located near the end of this prosodic domain.

Of course, there will be a certain amount of unsystematic variation left in the timing of tonal targets just as in other phonetic measures, like vowel targets and VOTs of plosive consonants. The number of possible realizations of intonation structures such as those in (2) are infinite in the same way as the number of realizations of the segmental structure.

Leaving aside biological or physical factors such as speaker gender or age, factors that may affect the phonetic implementation of tones and hence the acoustic signal produced, could be divided into two categories. Realizational differences can be driven by variation in the *phonological* context and the *paralinguistic* context in which the structure is pronounced. Variation of the first type is illustrated in (2) and (3), from Ladd (2008:45f), while (4) illustrates paralinguistic variation in intonation.

The falling-rising melodies in (3a) and (3b) can be used for echo questions in English. The  $f_0$  contours in the two examples are clearly different, yet they can be said to be instances of the same tonal morpheme. The difference in shape is a consequence of the length of the utterance that the tune is produced on. In (4a) and (4b), the rising melodies on *moving* differ in shape, as a consequence of the degree of surprise with which the questions are realized. In both examples, the

---

<sup>7</sup> Following Gussenhoven (2004), I assume that tones may have double alignment, which means they will have two locations in the segmental string. The trailing tone L of a prenuclear accent H\*L, for example, may align with the right edge of the starred tone to its left (H\*) and the left edge of the starred tone to its right (e.g., H\*), creating a flood plain in the realization.

tunes of the (a) and (b) examples are phonologically identical but phonetically different.

- (3) a. A: *I hear Sue's taking a course to become a driving instructor.*



B: **SUE!?**

- b. A: *I hear Sue's taking a course to become a driving instructor.*



B: **A DRIVING instructor!?**

- (4) a. A: *Do you have some spare boxes? I'm moving out today.*



B: **You're MOVing?**

(I knew you were going to move out but not that it was today.)

- b. A: *Do you have some spare boxes? I'm moving out today.*



B: **You're MOVing?**

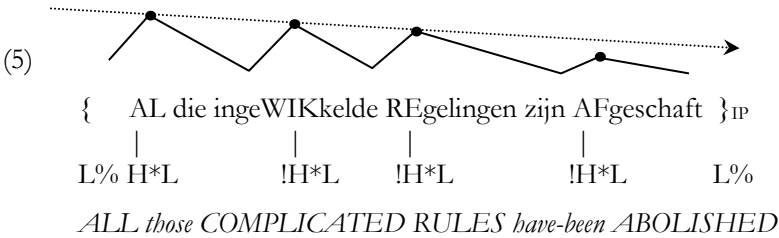
(I hadn't the faintest idea that you were planning on moving.)

The experiments described in this dissertation were designed to investigate realizational variation between varieties of Dutch that stem from a) differences in phonological context and b) differences in paralinguistic context. Below, I will provide examples of phonologically-driven variation in the realization of intonation structures, which can be captured by a set of *phonetic implementation rules*. It explains how various tonal (2.2.1), segmental (2.2.2) and prosodic factors (2.2.3) are involved in 'fine-tuning' the timing and scaling of tonal targets. Next, 2.2.4 briefly introduces the place of paralinguistics in intonation and give examples of how paralinguistic factors may affect the realization of intonation structures.

### 2.2.1 Tonal context

The tonal context in which a tone occurs will affect the scaling of its target, as observed above. Different H-tones within an intonation contour will be realized at different pitch heights as a consequence of phonetic implementation rules such as Downstep, Final Lowering or Upstep.

Downstep is a stepwise lowering of tonal targets, and is common in African languages with lexical tone. Pierrehumbert (1980) was the first to propose a phonetic implementation rule of Downstep to account for scaling variation in an intonation-only language, American English. In this dissertation, I follow Gussenhoven (1991) and van den Berg, Gussenhoven, and Rietveld (1992), who analyze Downstep in English and Dutch as an optional morpheme which downsteps  $H^*$  after  $H(L)$  <sup>8</sup>. Its effect is illustrated in (5) for Standard Dutch, adapted from Gussenhoven (2005) and originally from Collier and 't Hart (1981). In other languages, Downstep may be triggered by a different tonal context and may have different realizational effects.



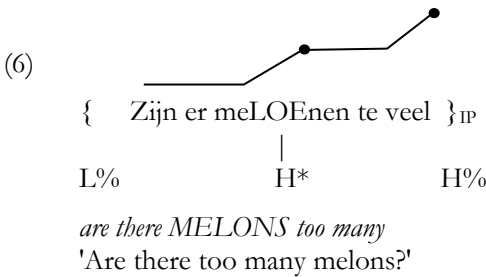
The downstepped high tone is notated as  $!H$ . Crucially, the phonological status of the tone as  $H$  is not altered by Downstep. Without the phonetic implementation rule, an analysis of the pitch peaks in (5) would require four different tonal categories, whereas now all peaks can be transcribed with  $H$ .

Final Lowering is a rule proposed by Liberman and Pierrehumbert (1984) to account for the fact that in a series of Downstepped pitch accents, the last accent is scaled lower than would be predicted by application of a constant lowering factor. This is visible in (5), where the pitch peak on the final accented syllable is lower than predicted by the arrow.

Finally, Upstep is a rule that raises the target of a high final boundary tone ( $H\%$ ) after an immediately preceding  $H\%$ . An example of Upstep is given in (6) for Standard Dutch (from Gussenhoven 2005).

<sup>8</sup> The Downstep morpheme is assumed to apply to all non-initial pitch accents in the utterance. The final pitch accent may be exempt from being downstepped.

<sup>9</sup> In her analysis of American English, Pierrehumbert (1980) proposed Upstep to apply to  $L\%$  after a phrase accent  $H$ - to account for level pitch at the end of an utterance. In the ToDI-style analysis adopted in this dissertation, phrase-final level pitch is accounted for by the absence of a boundary tone after  $H^*$  or  $L^*H$ , and thus doesn't require a phonetic implementation rule (Gussenhoven 2004:299).



### 2.2.2 Segmental effects on timing

Besides the tonal context, the segmental composition and CV-structure of syllables may also affect the timing of tonal targets<sup>10</sup>. Such effects have been reported for a variety of languages, including Dutch.

In an experiment that investigated the perception of Dutch final falls as early or late (i.e., downstepped or non-downstepped), Rietveld and Gussenhoven (1995) found that the timing of the nuclear peak is affected by the CV-structure and segmental composition of the accented syllable. The perceived pitch peak is located earlier, relative to the beginning of the accented vowel, the more consonants are added to the syllable onset. Moreover, voiced onset consonants exert an additional leftward pull on the timing of the peak, whereas voicing in the coda pulls the peak to the right. Rietveld and Gussenhoven showed that listeners are sensitive to these timing differences, because their perception of early and late peaks as downstepped or non-downstepped accents is different, depending on segmental structure. The authors account for the structure-dependent timing of the peak by assuming that it facilitates a 'comfortable realization of the pitch features concerned' (Rietveld and Gussenhoven 1995:383).

For Spanish, Prieto and Torreira (2007) similarly found that the timing of the accentual peak in LH\* prenuclear accents depends on the presence of a coda consonant. Peaks were found to be located at the end of the accented vowel in CV syllables, but around the beginning of the sonorant coda in CVC syllables. This effect was independent of the syllable's duration, which is known to correlate with the timing of tonal targets (cf. Silverman and Pierrehumbert 1990, Prieto, van Santen, and Hirschberg 1995). As Prieto and Torreira mention in their introduction, there are a number of other languages, besides Spanish and Dutch, in which peak timing depends on syllable structure, including several

<sup>10</sup> This section does not describe microprosodic effects on f<sub>0</sub> such as higher intrinsic f<sub>0</sub> on high vowels, because such segmental effects are usually ignored by the human ear. See Gussenhoven (2004:5-10).

varieties of Italian (D’Imperio 2000, Gili-Fivela and Savino 2003), English (van Santen and Hirschberg 1994), French (Welby and Loevenbruck 2005, 2006) and German (Mücke et al. 2009).

Ladd, Mennen and Schepman (2000), and Schepman, Lickley and Ladd (2006) show that the timing of prenuclear and nuclear peaks in Dutch is affected by vowel quality. In prenuclear accents, the peak on accented syllables with the tense vowel /i/ occurred at the end of the vowel, whereas for the lax vowel /ɪ/, the peak occurred late in the following consonant. In nuclear accents, the peak also occurred later if the accented vowel was lax, although still within the vowel<sup>11</sup>.

### 2.2.3 Prosodic context

A third set of phonetic implementation rules captures variation that arises from the prosodic context in which intonation structures are pronounced. A number of studies have shown that an upcoming prosodic boundary such as a syllable, word or IP boundary, or an upcoming (nuclear or prenuclear) accented syllable may affect the timing of tones. Generally, the smaller the distance to the next boundary, the earlier the location of the tonal target, mostly the pitch accent peak<sup>12</sup>.

In 2.2.2, we have already seen for Spanish that peak timing is earlier if the distance of the accented vowel to the next syllable boundary is shorter (e.g., in CV syllables vs. CVC syllables (Prieto and Torreira 2007)).

Silverman and Pierrehumbert (1990) demonstrated this ‘peak retraction’ effect for English, where peak timing is earlier in contexts where the distance of the accented syllable to the next word boundary is smaller (measured in number of unstressed syllables). They showed that the prenuclear peak in names like *Mom*<sub>nb</sub> *Le Mann* occurred earlier than in *Mama*<sub>nb</sub> *Lemm*. The effect has been replicated in other languages, including Mexican Spanish (Prieto, van Santen, and Hirschberg 1995), and Central Catalan and Peninsular Spanish (e.g., Prieto, Estebas, and Vanrell 2010).

---

<sup>11</sup> Even though phonological properties of the syllable may play a role in the timing of tonal targets, it is not the case that individual tones are phonologically ‘associated’ to elements of the segmental string, as has been suggested by the term ‘segmental anchoring’ in literature of the last fifteen years. Following Ladd (2004:127-128, 2008:172-180), I consider segmental anchoring to be a phonetic observation that could be empirically tested in other languages. It is a property of *some* tones in *some* contexts and in *some* languages, and is as such not a ‘basic’ or default characteristic of all tones.

<sup>12</sup> Many studies investigating the timing of targets have found that the location of the low tone preceding the peak is often unaffected by prosodic context, being anchored stably to the onset of the accented syllable (e.g., Caspers and van Heuven 1993, Arvaniti, Ladd, and Mennen 1998, Ladd et al. 1999).



Thirdly, peak timing is known to be affected by the distance of the accented syllable to the next accented syllable (in number of unaccented syllables). Silverman and Pierrehumbert (1990) show this for English, where the peak of the pitch accent is earlier in *MAma LEMM* than in *MAma Le MANN*. In 'accent clash' situations, such as *MOM LEMM*, with two adjacent accented syllables, the peak is earliest<sup>13</sup>.

The next prosodic boundary to give a peak retraction effect is the IP boundary: peaks tend to be located earlier in the segmental string, the closer the accented syllable is to the right edge of the IP (e.g., Steele 1986, Prieto, van Santen, and Hirschberg 1995, Dalton and Ní Chasaide 2007, Ladd et al. 2009).

A number of studies have demonstrated that peaks in nuclear pitch accents are aligned earlier than peaks in otherwise comparable prenuclear pitch accents (Silverman and Pierrehumbert 1990, Prieto, van Santen, and Hirschberg 1995, Schepman, Lickley, and Ladd 2006, Dalton and Ní Chasaide 2007a, Ladd et al. 2009, Mücke et al. 2009). The difference is mostly accounted for by the assumption that there is more pressure on the realization of nuclear tones, due to the presence of extra tones such as boundary tones or a phrase accent (cf. Silverman and Pierrehumbert 1990, Hualde et al. 2002, Ladd 2008:142, Ladd et al. 2009), in other words a consequence of *tonal crowding*.

As Schepman, Lickley, and Ladd (2006) point out, tonal crowding may arise from structural factors such as the number of unstressed syllables between two accented syllables, but also from more phonetic factors such as the actual distance (in ms) between the two accented syllables. The structural distance between the accented syllables in *BLOOmington VALley*, for example, is two unstressed syllables. While this is the same for *TRInity COLLege*, the phonetic distance is smaller in the second example, due to shorter durations of the segmental material. This distance becomes even smaller if the noun phrase is realized at a higher speech rate. Both phonetic distance and speech rate may affect the timing of targets in the same way as structural factors such as the right-hand prosodic context. Generally, the shorter the accented syllable, the earlier the peak (e.g., Silverman and Pierrehumbert 1990, Caspers and van Heuven 1993, Prieto, van Santen, and Hirschberg 1995). Caspers and van Heuven (1993) also report scaling effects for Dutch, showing that under time pressure arising from a higher speech rate, the entire pitch contour is raised (i.e., a change in *pitch register*), without affecting its excursion.

---

<sup>13</sup> The original term for this context is 'stress clash' in Silverman and Pierrehumbert (1990:78), but in fact this is a case of accent clash. Schepman, Lickley, and Ladd (2006) investigated for Dutch whether a stress clash situation, whereby a nuclear accented syllable is immediately followed by a stressed but unaccented postnuclear syllable, has similar effects as an 'accent clash' situation (compare *BLUEberry* and *BLUE BERRY*) but found no evidence for peak retraction in this context.

Considering that effects on tonal timing are stronger when the prosodic boundary is ranked higher (see Prieto, Estebas, and Vanrell 2010), and when tones are closer together, one can expect that time pressure gets particularly high in IP-final position. Indeed, we see that the realization of an IP-final intonation contour, where the time available for its execution is limited, can be rather different from a non-final contour, and that languages have different ways of dealing with such pressure.

#### **2.2.4 Pitch variation for paralinguistic purposes<sup>14</sup>**

On a parallel channel of communication, speakers use pitch variation<sup>15</sup> to signal paralinguistic meaning, which involves the speaker's attitude towards the message (informational meaning, e.g. 'assertive', 'emphatic') and towards the listener (affective meaning, e.g., 'polite', 'aggressive', 'surprised'). Such paralinguistic meanings would appear to be derived from anatomical and physiological effects on vocal fold vibration (Gussenhoven 2002). The Frequency Code (Ohala 1983) reflects the correlation between the size of the larynx and the vibration of the vocal folds, which is faster for smaller vocal folds, hence producing higher pitch. Although a speaker's larynx size is obviously anatomically determined, he or she can manipulate pitch height to express the power relations that are associated with smaller and larger vocal folds. Higher pitch tends to be associated with femininity, submissiveness, or uncertainty, whereas low pitch expresses masculinity, dominance, or certainty. In similar vein, Gussenhoven (2002) proposed two more 'codes'. The Production (Phase) Code is related to energy phasing. The decline in subglottal air pressure from the beginning of the exhalation phase to the end correlates with the gradual lowering of  $f_0$  throughout the utterance, also known as declination. The beginning of an utterance may consequently be associated with high pitch, whereas the end is associated with low pitch. Using the Production Code, speakers may vary the pitch at the beginning of utterances to signal informational meanings like newness of topic (high initial pitch) or continuation of topic (low initial pitch) and at the end of utterances to signal continuation (high final pitch) or finality (low final pitch). Finally, the Effort Code reflects the energy level, and hence the articulatory precision (de Jong 1995), with which the utterance is pronounced. An utterance that is produced with more energy will have a wider pitch excursion.

---

<sup>14</sup> I refer the interested reader to Gussenhoven (2004, chapters 4 and 5) and Chen (2005) for a more elaborate introduction to the topic of paralinguistic intonation.

<sup>15</sup> Pitch variation forms part of a larger set of cues to signal paralinguistic meaning, also including variation in loudness, duration and voice quality, the use of particles and grammatical categories such as pejoratives, and visual cues such as a facial expression and body language. See, e.g., Gussenhoven (2004:24) and Ladd (2008:38).

This cue may be exploited by the speaker to signal enthusiasm, emphasis, or the informational meaning of ‘significant’. When the information weight of the communication is increased, the pronunciation of the pitch accent that expresses this may be enhanced. Besides large pitch excursions, this may also lead to, e.g., higher and later pitch peaks.

### 2.3 Cross-linguistic variation in intonation

The use of pitch varies to a greater or lesser extent in different languages. An obvious difference is between those languages that use pitch for lexical purposes and those that do not. But even within the class of intonation languages, variation is observed. Taking Dutch as the language of comparison, some languages are rather similar, like other West-Germanic languages such as English and German, whereas others differ in more respects, like the Romance language French. As explained earlier in Chapter 1.2, there are ‘different types of differences’ between languages, which may be categorized as (1) semantic, (2) systemic, (3) phonotactic and (4) realizational. Some differences may not in fact readily fall in any of these categories or fall into more than one. Indeed, Ladd (2008:116) adds a disclaimer, warning the reader that “no great theoretical significance should be attached to the taxonomy of differences”. Below, I will briefly define the four dimensions of variation and illustrate each with examples. As the main topic of this dissertation, realizational differences are given some special attention.

#### *Semantic variation*

Semantic variation concerns differences in the semantic meaning of phonologically identical tunes. Ladd (2008:117) illustrates this dimension with the calling contour, which is used in both North American English and German (and a range of other languages) to call another person, e.g. ‘*Da-vid!*’ with sustained high and mid level pitch on the first and second syllable, respectively. However, whereas the tune can be commonly used by an adult to call a child in both languages, the use between adults is restricted to German. Another example would be the use of the H\*L pitch accent, which may be used to signal both broad and corrective focus in English. In European Portuguese, the same pitch accent is used to express corrective focus, but cannot be used for broad focus, which is expressed by HL\* (Frota 2000). Similarly, whereas a rising intonation contour can be used to signal a question in most varieties of British English, speakers of Urban North British (UNB, particularly speakers from the Glasgow area) use the same rising intonation for statements. Other varieties of British use falling intonation in this case (Cruttenden 1986, see also Ladd 2008:127).

*Systemic variation*

Systemic variation occurs when, irrespective of semantic differences, language (varieties) have different inventories of phonologically distinct tune types. A case in point is French, which unlike English, does not have a low pitch accent (L\*) in its inventory (Post 2000 and Gussenhoven 2004:266). Second, in a comparative study on dialect intonation, Grabe et al. (2000) showed that Belfast English speakers use a different nuclear contour tune (a ‘rise-plateau’, or L\*H 0%) for statements and questions than speakers of three other varieties of English, who used falls and rises, respectively.

*Phonotactic variation*

Phonotactic variation is related to differences in the distribution and combination of boundary tones and pitch accents, and to differences in tune-text association. To take an example from French again, H\*L is a pitch accent of this language, which is however restricted to the use in prenuclear position, unlike in Dutch where it can occur in both prenuclear and nuclear position (Post 2000 and Gussenhoven 2004:267).

*Realizational variation*

Realizational differences between languages can be described along two dimensions. The first type of variation can be called *context-specific*, and occurs when languages respond differently to a particular context (like those described in section 2.2). The second type of variation may be called *language-specific* and describes realizational differences between languages or language varieties in more general terms. A well-known example is that British English speakers use a significantly wider pitch range than Dutch speakers (Willems 1982, de Pijper 1983). Similarly, van Bezooijen (1993, see also Gussenhoven 2004:81) found that Dutch has higher pitch if spoken by Belgian women than by Dutch women. Two language varieties may also differ in the realization of the same pitch accent in a ‘normal’ (e.g., non-final) context in terms of timing. Atterer and Ladd (2004) find that speakers of Southern German generally start prenuclear rises later than speakers of Northern German. The difference is robust but also small, which leads the authors to conclude that both rises belong to the same phonological category, but have different realizations, depending on the variety. Other language-specific variation involves, e.g., differences in segmental duration, pitch span or register.

Note that realizational differences also arise *within* languages. In section 2.2, we have already seen that numerous factors may affect the realization of pitch contours without changing their phonological identity. In addition, within-language realizational variation also occurs when a language uses the one phonological contour, but with different acoustic realizations, to express pragmatic differences such as statement versus questions. Under semantic variation above, I mentioned that speakers of Urban North British use a rising contour to express both statements and questions. However, the realization of

the rise is different in these contexts, with the rise in questions being higher than the rise in statements (Ladd 1996:144).

## 2.4 Variation: phonology or phonetics?

Even though this dissertation is mainly concerned with phonetic variation, it should be noted that there may not always be a sharp divide between phonetic and phonological variation. Imagine two dialects with a difference in peak timing. When should this be interpreted as a realizational difference of the same phonological category, as Atterer and Ladd (2004) did for Northern and Southern German (see realizational variation above) and when can we speak of two phonological contours that are distinguished by the alignment of the peak relative to the syllable, as in Frota (2000) on European Portuguese pitch accents for broad and corrective focus (see semantic variation above), or in D'Imperio and House (1997) on Neapolitan Italian rising and falling pitch accents for questions and statements? A traditional way of determining whether a difference in realization reflects a categorical distinction is by looking at data in a scatter plot. Two separate clouds of data points is taken to reflect two categories, whereas a continuum of data points reflects gradient variation. Ladd (2008:152) illustrates how difficult it can be to determine whether variation is phonetic or phonological by citing data from Braun (2006). Braun found that two pragmatic categories (contrastive vs. non-contrastive theme) showed significant differences in intonational realization, yet the data did not show two clear clusters of data points associated with contrast and non-contrast, respectively. Braun's results show that two distinct pragmatic categories may be realized by a continuum of phonetic realizations, whereby one end of the continuum is largely associated with one interpretation, and the other end with the opposite interpretation, but whereby the realizations in between are typically not clearly associated with one or the other.

Conversely, it may be the case that a continuum of phonetic realizations is interpreted as reflecting one or the other phonological category. When Dutch listeners were asked to interpret a continuum of early-to-late peak timing as either *low-sounding* (downstepped) or *high-sounding* (non-downstepped), their responses showed a clear S-shaped curve (Rietveld and Gussenhoven, 1995).

These examples show that we should be careful when interpreting variation as a reflection of two distinct categories as opposed to phonetic variants of a single category. There may be gradient variation whereby one end of the continuum is more closely associated with category A, and the other with category B.

## References

- Arvaniti, A., Ladd, D.R., and Mennen, I. (1998). Stability of tonal alignment: the case of Greek prenuclear accents. *Journal of Phonetics* 26, 3-25.
- Atterer, M. and Ladd, D.R. (2004). On the phonetics and phonology of “segmental anchoring” of F0: evidence from German. *Journal of Phonetics* 32, 177-197.
- Beckman, M.E., Hirschberg, J., and Shattuck-Huffnagel, S. (2005). The original ToBI system and the evolution of the ToBI framework. In: Jun, S-A (Ed.), *Prosodic typology: The phonology of intonation and phrasing*. Oxford University Press, 9-54.
- Braun, B. (2006). Phonetics and phonology of thematic contrast in German. *Language and Speech* 49, 451-493.
- Caspers, J. and van Heuven, V.J. (1993). Effects of time pressure on the phonetic realization of the Dutch accent-lending pitch rise and fall. *Phonetica*, 50, 161-171.
- Chen, A. (2005). *Universal and language-specific perception of paralinguistic intonational meaning*. PhD thesis. Utrecht: LOT.
- Collier, R. and 't Hart (1978). *Cursus Nederlandse intonatie*. Diepenbeek: Wetenschappelijk Onderwijs Limburg. Reprinted (with minor revisions) 1981, Louvain: Acco.
- Cruttenden, A. (1986). *Intonation*. Cambridge: Cambridge University Press.
- Dalton, M. and Ní Chasaide, A. (2007). Melodic alignment and micro-dialect variation in Connemara Irish. In: Gussenhoven, C. and Riad, T. (Eds.), *Tones and tunes. Volume 2: Experimental studies in word and sentence prosody* (pp. 293-315). Berlin: Mouton de Gruyter.
- Dalton, M. and Ní Chasaide, A. (2007a). Nuclear accents in four Irish (Gaelic) dialects. In: *Proceedings of the 16th International Conference of Phonetic Sciences (ICPhS)*, Saarbrücken, 965-968.
- de Jong, K.J. (1995). The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *Journal of the Acoustic Society of America* 97, 491-504.
- de Piiper, J. (1983). *Modelling British English intonation*. Dordrecht: Foris.
- D’Imperio, M. (2000). *The role of perception in defining tonal targets and their alignment*. PhD thesis, The Ohio State University.
- D’Imperio M. and House, D. (1997). Perception of questions and statements in Neapolitan Italian. In: Kokkinakis, G., Fakotakis, N., and Darmatas, E. (Eds.), *Proceedings of Eurospeech '97*, vol. 1. Rhodes. 1251-1254.
- Frota, S. (2000). Prosody and focus in European Portuguese. Phonological phrasing and intonation. PhD thesis, University of Lisbon. New York: Garland Publishing.
- Gili Fivela, B. and Savino, M. (2003). Segments, syllables and tonal alignment: a study on two varieties of Italian. In: *Proceedings of the 15th International Conference of Phonetic Sciences (ICPhS)*, Barcelona, 2933-2936.
- Goldsmith, J.A. (1976). *Autosegmental phonology*. PhD thesis, MIT. New York: Garland Press.
- Grabe, E. (1998). Pitch accent realization in English and German. *Journal of Phonetics*, 26, 129-144.
- Grabe, E., Post, B., Nolan, F., and Farrar, K. (2000). Pitch accent realization in four varieties of British English. *Journal of Phonetics*, 28, 161-185.

- Gussenhoven, C. (1991). Tone segments in the intonation of Dutch. In: Shannon, T.F. and Snapper, J.P. (Eds.), *The Berkeley Conference on Dutch Linguistics 1989* (pp. 139-155). Lanham, MA: University Press of America.
- Gussenhoven, C. (2002). Intonation and interpretation: phonetics and phonology. In: Bel, B. and Marlien, I. (Eds.), *Proceedings of Speech Prosody 2002, Aix-en-Provence*, 47-57.
- Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge: Cambridge University Press.
- Gussenhoven, C. (2005). Transcription of Dutch intonation. In: Jun, S.-A. (Ed.), *Prosodic typology: The phonology of intonation and phrasing*. Oxford: Oxford University Press, 118-145.
- Gussenhoven, C. (2016). The analysis of intonation: The case of MAE\_ToBI. *Laboratory Phonology Journal of the Association for Laboratory Phonology*, 7(1), 10.
- Gussenhoven, C., Rietveld, T., Kerkhoff, J., and Terken, J. (2002). *Transcription of Dutch Intonation* (2<sup>nd</sup> ed.). Available from: <http://todi.let.kun.nl/ToDI/home.htm>.
- He, X. (2012). *Mandarin-accented Dutch Prosody*. PhD thesis. Utrecht: LOT Publications.
- Horne, M. (Ed.) (2000). *Prosody: Theory and experiment. Studies presented to Gösta Bruce*. Dordrecht: Kluwer.
- Hualde, J., Elordieta, F., Gaminde, I., and Smiljanić, R. (2002). From pitch-accent to stress-accent in Basque. In: Gussenhoven, C. and Warner, N. (Eds.), *Laboratory Phonology 7* (pp. 547-584). Berlin and New York: Mouton de Gruyter.
- Ladd, D.R. (1996). *Intonational phonology*. Cambridge: Cambridge University Press.
- Ladd, D.R. (2008). *Intonational phonology* (2<sup>nd</sup> ed.). Cambridge: Cambridge University Press.
- Ladd, D.R., Mennen, I., and Schepman, A. (2000). Phonological conditioning of peak alignment in rising pitch accents in Dutch. *Journal of the Acoustical Society of America*, 107, 2685-2696.
- Ladd, D.R., Faulkner, D., Faulkner, H., and Schepman, A. (1999). Constant 'segmental anchoring' of F0 movements under changes in speech rate. *Journal of the Acoustical Society of America* 106, 1543-1554.
- Ladd, D.R., Schepman, A., White, L., Quarmby, L.M., and Stackhouse, R. (2009). Structural and dialectal effects of pitch peak alignment in two varieties of British English. *Journal of Phonetics*, 37, 145-161.
- Leben, W.R. (1973). *Suprasegmental phonology*. PhD thesis, MIT. Published 1980 by Garland Press, New York.
- Lieberman, M. and Pierrehumbert, J. (1984). Intonational invariance under changes in pitch range and length. In: Aronoff, M and Oehrle, R.T. (Eds.), *Language and sound structure*. Cambridge, MA and London: MIT Press, 157-233.
- Mücke, D., Grice, M., Becker, J., and Hermes, A. (2009). Sources of variation in tonal alignment: Evidence from acoustic and kinematic data. *Journal of Phonetics* 37, 321-338.
- O'Connor, J.D. and Arnold, G.F. (1973). *Intonation of colloquial English* (2<sup>nd</sup> ed.). London: Longman.
- Ohala, J. (1983). Cross-language use of pitch: an ethological view. *Phonetica* 40, 1-18.
- Pierrehumbert, J.B. (1980). *The phonetics and phonology of English intonation*. PhD thesis, MIT.
- Post, B. (2000). *Tonal and phrasal structures in French intonation*. The Hague: Holland Academic Graphics.

- Prieto, P., Estebas-Vilaplana, E., and Vanrell, M.M. (2010). The relevance of prosodic structure in tonal articulation. Edge effects at the prosodic word level in Catalan and Spanish. *Journal of Phonetics* 38, 687-705.
- Prieto, P. and Torreira, F. (2007). The segmental anchoring hypothesis revisited. Syllable structure and speech rate effects on peak timing in Spanish. *Journal of Phonetics* 35, 473-500.
- Prieto, P., van Santen, J., and Hirschberg, J. (1995). Tonal alignment patterns in Spanish. *Journal of Phonetics*, 23, 429-451.
- Rietveld, T. and Gussenhoven, C. (1995). Aligning pitch targets in speech synthesis: effects of syllable structure. *Journal of Phonetics* 23, 375-385.
- Schepman, A., Lickley, R., and Ladd, D.R. (2006). Effects of vowel length and 'right context' on the alignment of Dutch nuclear accents. *Journal of Phonetics* 34, 1-28.
- Silverman, K. and Pierrehumbert, J.B. (1990). The timing of prenuclear high accents in English. In J. Kingston and M. Beckman (Eds.), *Papers in laboratory phonology I* (pp. 72-106). Cambridge: Cambridge University Press.
- Steele, S. (1986). Nuclear accent f0 peak location: Effect of rate, vowel, and number of syllables. *Journal of the Acoustical Society of America*, 80 (Suppl. 1), s51.
- 't Hart, J., Collier, R., and Cohen, A. (1990). *A Perceptual Study of Intonation: An Experimental-Phonetic Approach to Speech Melody*. Cambridge: Cambridge University Press.
- van Bezooijen, R. (1993). Verschillen in toonhoogte: Natuur of cultuur? *Gramma/TTT* 2, 165-179.
- van den Berg, R., Gussenhoven, C., and Rietveld, T. (1992). Downstep in Dutch: Implications for a model. In G. Docherty and D.R. Ladd (Eds.), *Papers in Laboratory Phonology II: Gesture, Segment, Prosody* (pp. 335-359). Cambridge: Cambridge University Press.
- van Santen, J.P.H. and Hirschberg, J. (1994). Segmental effects on timing and height of pitch contours. In: *Proceedings of the International Conference on Spoken Language Processing 1994*, 719-722.
- Welby, P. and Lævenbruck, H. (2005). Segmental "anchorage" and the French late rise. In: *Proceedings of INTERSPEECH 2005, Lisbon*, 2369-2372.
- Welby, P. and Lævenbruck, H. (2006). Anchored down in Anchorage: Syllable structure, rate, and segmental anchoring in French. *Rivista di Linguistica* 18(1), 74-124.
- Willems, N. (1982). *English intonation from a Dutch point of view: An experimental-phonetic investigation of English intonation produced by Dutch native speakers*. H.-I. Ambacht: Intercontinental Graphics.



# REGIONAL VARIATION IN PHONETIC RESPONSES TO TIME PRESSURE IN DUTCH IP-FINAL NUCLEAR CONTOURS<sup>1</sup>

---

## Chapter 3

### Abstract

We conducted a production experiment with 119 speakers of six regional varieties of Dutch, who realized nuclear falls, rises, and fall-rises on IP-final monosyllables. In each contour condition, the duration of the sonorant rime of the nuclear accented word was varied in order to test the effect of the distance of the upcoming intonational phrase boundary on the phonetic realization of the contour. We found that the fall-rise was used less as the duration of the sonorant rime was shorter. In the phonetic domain, the range of adjustments fell into three broad categories. First, speakers of some varieties *took shortcuts* by truncating final pitch movements and by undershooting the low target of the fall-rise. Second, speakers *worked harder* by compressing contours. Third, speakers appeared to *take more time* by increasing the duration of the rime in the fall-rise condition, while also retracting the peak of the fall-rise contour. We conclude that a description in terms of only *truncation* and *compression* is inadequate for our data. In line with earlier investigations, the responses were shown to be dialect-specific and contour-specific. Moreover, in the cases of rises they were shown to be dependent on the segmental composition of the accented syllable.

### 3.1 Introduction

What is functionally the same intonation contour may phonetically be very different depending on the context in which it occurs. One of the factors influencing the realization of contours is the time available for its execution. Since  $f_0$  movements require a minimum amount of time to be fully executed (Xu and Sun 2002), the production of targets can be compromised under tonal crowding (Pierrehumbert 2000). The magnitude of time pressure in intonation-only languages like Dutch and English depends, first, on the available space between tonal targets, for example between a prenuclear and nuclear pitch accent, or between a nuclear pitch accent and a phrasal boundary. There is less time

---

<sup>1</sup> This chapter is a slightly revised version of: Hanssen, J., Peters, J., and Gussenhoven C. (ms). Regional variation in phonetic responses to time pressure in Dutch IP-final nuclear contours. Submitted to *Journal of Phonetics*.

available for the execution of the prenuclear and nuclear accents in *a skittish kitten*, which has shorter vowels and less voicing as compared to *a roaring lion*, with identical syllable structure but more voicing. The available time in *black cat* is, in turn, less than in *a skittish kitten*, which follows from a lack of intervening unaccented syllables ('accent clash' Silverman and Pierrehumbert 1990) and a shorter space between the nuclear accent and the intonational phrase edge (e.g., Steele 1986). The shorter the stretch of sonorant segmental material there, the more the realization of nuclear and boundary tones will be under pressure (e.g., Grabe 1998b). There will obviously be more pressure on the production of contours that consist of three movements, such as the rise to H\*, fall to L and final boundary rise to H% of the complex falling-rising tune, than on those that consist of only two, such as a falling contour. Also, tonal crowding will be more problematic in fast speech than in normal or slow speech, because the time to produce tonal targets is reduced due to shorter segmental durations (Caspers and van Heuven 1993 for Dutch).

This paper reports what strategies speakers adopt in response to time pressure due to (1) limited sonorant material in IP-final position and (2) to complexity of the IP-final nuclear contour. More specifically, we show how speakers of five Dutch regiolects and Standard Dutch accommodate falls (H\*L L%), rises (L\*H H%, L\* H%, H\* H%) and fall-rises (H\*L H%) on four IP-final monosyllabic words with decreasing duration of the sonorant portion in them. We make a distinction between phonological and phonetic responses, and place the latter in one of three categories. Speakers can (i) take shortcuts, (ii) work faster, and (iii) take more time. Within the first category, speakers can choose to truncate movements or undershoot targets, the second category consists of compression, and the third category comprises timing adjustments such as peak retraction, and durational adjustments in the segmental domain.

Traditionally, f0 adjustments in response to IP-final time pressure have been interpreted in terms of two economizing mechanisms. The first is "truncation", whereby speakers produce an incomplete version of the contour, leaving the speed of the f0 movement intact. The second is "compression", whereby the full contour is produced, but at a higher speed. The terms truncation and compression<sup>2</sup> were originally introduced by Erikson and Alstermark (1972) and Bannert and Bredvad (1975) to describe speakers' responses to varying voicing durations of non-final accent 1 and accent 2 stressed syllables in varieties of Swedish. They found that the effects of shortening the stressed vowel on the f0 contour depend on dialect and word accent; speakers could either truncate or

---

<sup>2</sup> The mechanism of *compression* was originally introduced as *rate adjustment* in Erikson and Alstermark (1972) but renamed *compression* by Bannert and Bredvad (1975).

compress both accents, or truncate one accent and compress the other. Also, Grønnum (1989) reports for varieties of Danish that the effect of shortening the stressed vowel of “prosodic stress groups” (a stressed syllable plus any following unstressed syllables) depends on their position in the IP. In final stress groups, the  $f_0$  movement on the stressed syllable was truncated in all varieties, but in non-final stress groups, compression could be observed for some varieties. Grønnum’s (1989) findings for Northern German provided some initial evidence that German speakers truncate falling  $f_0$  movements, but that rising movements are compressed, suggesting a dependency on pitch accent type.

The effect of voicing duration on the shape of IP-final  $f_0$  contours was more systematically investigated by Grabe (1998a,b) for Northern German (Braunschweig) and Southern British English. Speakers of both varieties read short sentences with three experimental words in IP-final position (*Schiff*, *Schief*, and *Schiefer* for German, and *Shift*, *Sheaf*, and *Sheafer* for English). To interpret the effect of voicing duration on the  $f_0$  contours of the target words, Grabe measured “rate of  $f_0$  change”, which was calculated by dividing the maximum  $f_0$  excursion of the fall or the rise by the duration of that pitch movement. Comparing the shorter words to the longer words, an increase of the rate of  $f_0$  change was taken to reflect compression, whereas a stable rate of  $f_0$  change was considered to reflect truncation. The results showed that responses to time pressure are language-specific, and that moreover, the Northern German response is pitch accent-dependent. English compresses both falling and rising  $f_0$  movements, whereas German compresses rises, but truncates falls. The latter result is in line with Grønnum’s earlier results.

Peters (2006:162ff.) and Grabe et al. (2000) provide additional evidence that truncating and compressing responses to variation in voicing duration differ not only between languages, but also between varieties of the same language. Comparing the realization of IP-final falls on syllables with short and long vowels in Berlin and Hamburg German, Peters found that the Berlin speakers behave like Braunschweig German and truncate falls, but Hamburg speakers tend to compress falls on syllables with short vowels. More findings on compression and truncation in urban vernaculars of German are reported by Gilles 2005: 215ff. Grabe et al. (2000) found similar region-dependent responses to IP-final time pressure in their comparison of four varieties of English. They observed compression of falling and rising contours in Cambridge English and Newcastle English and truncation in Leeds English, while in Belfast English, in which speakers produced a rise for both experimental sentence types, speakers also tended to truncate.

There are two problems with the interpretation of IP-final time pressure responses in terms of *truncation* and *compression*. First, the use of a single phonetic measure (rate of  $f_0$  change) to reflect either truncation or compression is problematic. An increase in rate of  $f_0$  change is taken to reflect compression,

whereas a stable rate of  $f_0$  change is interpreted as truncation. However, an increase in the speed of the  $f_0$  movement may occur simultaneously with a reduction of the contour's  $f_0$  excursion. Speakers of Cambridge and Newcastle English increased the rate of  $f_0$  change on shorter words, but also reduced the  $f_0$  excursion, thereby showing some truncation in addition to compression (Grabe et al. 2000)<sup>3</sup>. Similarly, the results of a pilot investigation of time pressure effects in Standard Dutch suggest that steeper falling and rising slopes in shorter words also have smaller  $f_0$  excursions (Hanssen et al. 2007). Rate of  $f_0$  change by itself cannot capture such differences, a problem that becomes more acute when complex movements are considered. In a falling-rising tune, adjustment responses to time pressure, such as steeper slopes or smaller  $f_0$  excursions, can take place in the initial falling movement, in the final rising movement, or in both movements. Hanssen et al. (2007) showed that the source of decreasing rates of  $f_0$  change in shorter words could be found in both movements of the fall-rise, which were truncated and realized at slower speed.

Second, we can easily think of other responses to IP-final time pressure besides truncation and compression, on the basis of adjustment strategies that have been reported in the literature. A phonological response is to avoid complex contours in case of time pressure and replace them with simple ones. Phonetic responses other than truncation and compression are to undershoot targets, to change the timing of pitch accents, or to increase the segmental material on which the movement in realized.

As for the phonological adjustment, Leben (1976:81) notes that the L\*H LH% melody of American English, which is readily usable in phrases with the accented syllable in antepenultimate position, is avoided on final accented syllables. Similarly, Féry (1993:91) reports that the use of the fall-rise on a single syllable is somewhat marked in German. This observation is supported by Ladd (2008:183f), who claims that whereas it would be common for a speaker of German to use a falling-rising melody (H\*L H%) to ask 'Ist das IHR Geld?' (Is that YOUR money?), he would avoid the contour when he asks 'Ist das Ihr GELD?' (Is that your MONEY?), in which case a rise (H\* H%) is preferred. For tone languages, Zhang (2000) shows how complex tone structures are avoided on short-vowelled syllables. Finally, a falling-rising melody is avoided in the Limburgish dialect of Venlo on phrase-final syllables in the accent 1 tone class, but readily used for such syllables if they have accent 2 (Gussenhoven and van der Vliet 1999). Since final syllables with accent 2 are longer than those with accent 1 regardless of intonation contour, here too the explanation lies in the conflict between complexity of the pitch movement and available time. Instead of absolute avoidance, the frequency of use of complex contours may be less

---

<sup>3</sup> Grabe et al. (2000:174) argue that speakers respond to time pressure by compression of the fall in these varieties, and that the variation in  $f_0$  excursion is an effect of syllabic structure.

than expected. For Dutch, Lickley, Schepman, and Ladd (2005) claim that the fall-rise is dispreferred in questions when the nuclear accented syllable is IP-final. Results from a Map Task experiment show that in IP-final position, the proportion of questions realized with a fall-rise is 5% ( $n = 1$ ) on words with word-final stress, compared to 53% ( $n = 21$ ) and 61% ( $n = 19$ ) on words with penultimate and antepenultimate stress, respectively. Finally, different varieties may show different responses to the same situation. In British English, the use of the rise-fall-rise on ultimate stressed syllables such as *A car!?* may be executed in full. Similarly, the IP-final use of the fall-rise is observed in some regional varieties of German, such as Hamburgian (Gilles 2005:293, Peters 2006:91,164) and Mannheim (Gilles 2005:304).

To return to phonetic responses to time pressure, a third adjustment strategy is to undershoot targets. Gilles (2005:304) mentions that in Mannheim German, the pitch height of the low tone in fall-rises correlates with the number of syllables in the accented word. Words with fewer syllables have higher low tones, resulting in a particularly flat contour on monosyllables. Mannheim German is different from Hamburg German, which compresses fall-rises in IP-final position (Gilles 2005:94). We define ‘undershoot’ as a less extreme  $f_0$  value for a non-final target. In this interpretation, truncation is a special type of undershoot, that of a final target. Figure 1 depicts the three phonetic responses schematically for falls, rises and fall-rises. The time pressure is indicated by the vertical interrupted line, which serves as the end of the shortened sonorant portion. Adjusted contours are shown with black interrupted lines, and the unadjusted contours are given as gray uninterrupted lines. Truncation (first column) leaves the contour unaffected except for the non-realization of the final portion. Undershoot (second column) raises or lowers a low or high non-final target respectively; in the case of the fall-rise there are two options, as shown. Compression is shown in the third column.

A fourth phonetic response to time pressure is to realize targets earlier. In an investigation of the timing of nuclear peaks ( $H^*$ ) in American English, Steele (1986) observes that decreasing the distance between the accented syllable and the IP-final boundary (either by increased speech rate or by a reduction of the number of postaccentual syllables) caused peaks to align earlier, relative to the vowel onset of the accented syllable. Peak retraction was also found in Mexican Spanish (Prieto et al. 1995) and in the urban vernaculars of Berlin and Hamburg (Peters 1999), when the distance between the nuclear accented syllable and the end of the IP was reduced.

Finally, speakers can respond to time pressure by increasing the duration of the available sonorant material, so as to create more time for the execution of the pitch movement. Speakers may add duration when a movement is physiologically more difficult to produce, as in the case of rises (Ohala and Ewan 1973). They may also add duration if more movements, must be produced. Since fall-rises

involve both a downward and an upward pitch change, it is conceivable that speakers will lengthen IP-final accented syllables that carry a complex contour. Indeed, Gilles (2005:294) mentions that monosyllables in Hamburg are longer when carrying a fall-rise.

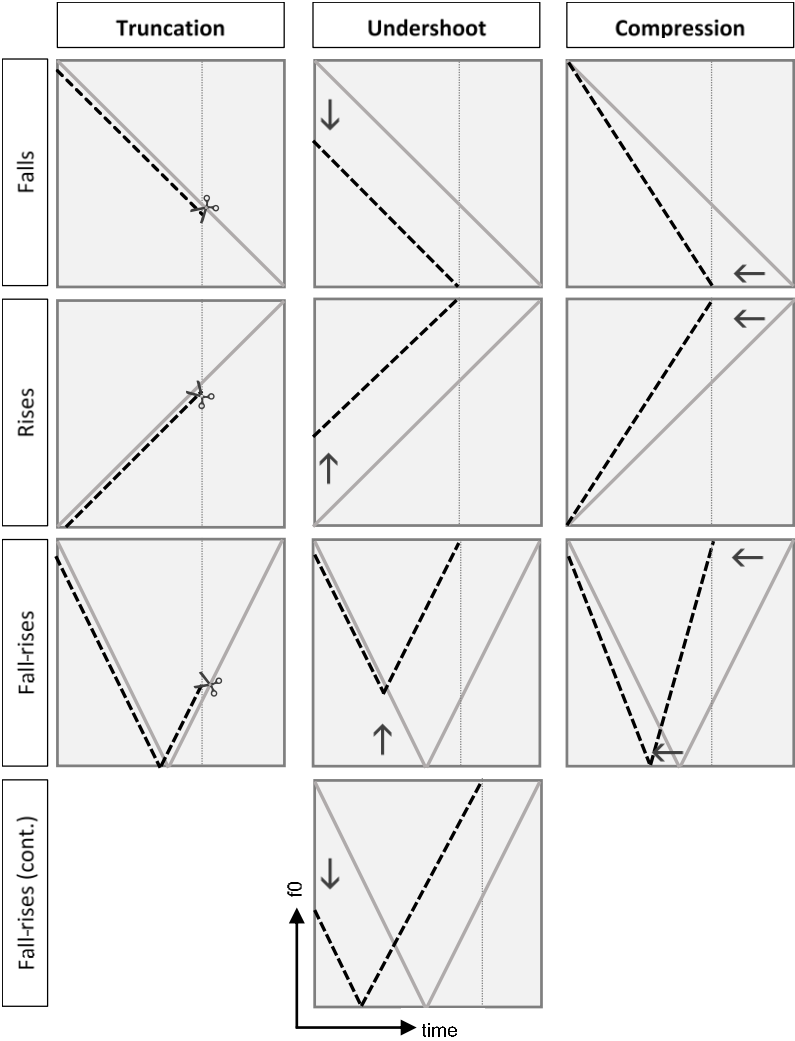


Figure 1. Schematic illustrations of truncation, undershoot and compression (columns) for falls, rises and fall-rises (rows). The gray uninterrupted lines represent an f0 contour on a long sonorant portion, while the black interrupted lines represent an f0 contour on a short sonorant portion. The arrows and scissors illustrate how the f0 contours have been affected by time pressure. Undershoot can affect a fall-rise in two ways: it affects either the dip between two high tones or the initial high tone.

We have already seen that speakers can combine truncating and compressing strategies within a single f0 contour. Speakers might in fact combine other phonetic adjustments. If truncation and undershoot apply simultaneously, both final targets and non-final targets are affected. Similarly, either or both truncation and undershoot can combine with compression, affecting both the scaling of targets and the speed of the f0 movement. These distinctions have not always been made in this way. For instance, when Grabe et al. (2000) measured the f0 excursion relative to f0 duration in *Shift*, *Sheaf*, and *Sheaffer*, they were able to show that truncation and compression applied at the same time. However, by itself, f0 excursion does not capture differences in scaling of either the higher target(s) or the lower target(s). That is, the position of the movement in the speaker's total pitch span is left unspecified.

To summarize, speakers have several options when confronted with insufficient time to produce f0 movements. They may replace f0 contours with simpler ones (a phonological response) or they may respond by altering the phonetic realization of the contour. Such phonetic adjustment may involve expending less effort in the production of contours by taking shortcuts, for example by truncating f0 movements and undershooting targets. Second, speakers may choose to work harder, producing full f0 movements in less time (compression). Third, speakers may take more time to produce a contour, by placing targets earlier or by increasing the duration of segments. Importantly, speakers may be expected to apply several pitch accommodation strategies simultaneously. The choice for a particular (set of) responses appears to depend on language or variety, and on pitch accent type.

### 3.1.1 Measuring alignment

Alignment differences are expressed either relative to a segmental reference point or as a proportion of a segmentally defined interval, as illustrated in Figure, 2, which schematically represents a falling pitch accent in a longer and shorter syllable rime. The f0 peak in short rimes (black contours) is retracted relative to the beginning of the syllable rime both in panel (a) and in panel (b). In panel (a), however, the speaker has also compressed the f0 contour to fit the syllable rime, while in panel (b) he has moved the contour left, keeping the f0 duration of the fall intact. While the peaks in panel (a) have different alignments in absolute terms, they actually both occur halfway through the rime in both the long and the short syllable, if measured as a proportion of that syllable. In panel (b), the peak occurs earlier in the short syllable than in the long syllable, regardless of whether we measure it with reference to the beginning of the rime or as a proportion of the syllable rime. The situation in panel (b) illustrates further that the choice of the absolute reference point will determine the result that is reported, since with

reference to the end of the rime, the peak alignments are the same (50 ms in both contours).

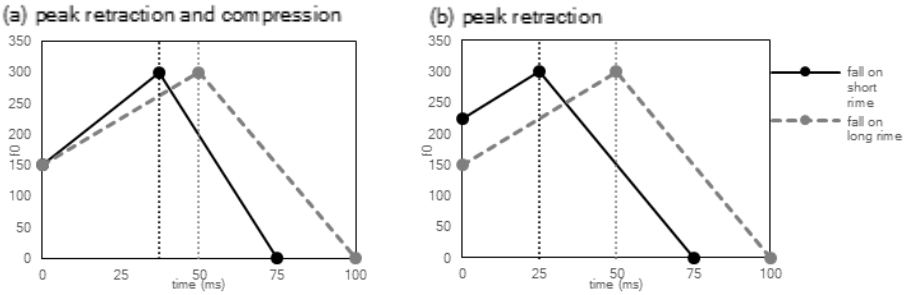


Figure 2. Schematic illustrations of falls on longer (gray interrupted lines) and shorter syllable rimes (black uninterrupted lines). The vertical reference lines mark the timing of the peak.

Given these different ways of measuring target alignment, we need to be explicit not only on how we measure it, but also on why we prefer one method over another. We will generally report alignments as a proportion of the sonorant rime duration. If shorter sonorant rime durations leave the relative alignment of the target unchanged, we will not consider this a case of target retraction, even though the absolute time from the beginning of the rime has decreased. The motivation for this is that the adjustment could be described as the regular placement of the pitch contour within the available sonorant space. If the proportional alignment in a shorter sonorant rime decreases, meaning that the period of time between the rime beginning and the target represents a smaller proportion of the total sonorant rime, we will report this additional adjustment as target retraction. In addition, we will report absolute alignments relative to the rime beginning where this seems relevant.

### 3.1.2 Phonetics or phonology?

Interpreting  $f_0$  variations as categorical or gradient is not always straightforward. For example, rather than as a phonetic process of curtailing the final part of a  $f_0$  contour, truncation could alternatively be interpreted as tone deletion (cf. Ladd 2008:183). Grabe investigated both options for her German and English data and concluded that both compression and truncation are gradient phonetic processes and that truncation is not the result of tone deletion. And whereas Ladd (2008:183f) presents the replacement of the fall-rise with the high rise as a phonological adjustment strategy, Peters (2006:91) suggests that a phonetic view of the same fact could be taken. In cases such as these, a perception experiment may provide the answer. In this way, Odé (2005) showed that Russian LH\*L and



LH\* accents on ultimate stressed syllables were realized almost identically due to truncation, not due to tone deletion or neutralization.

### 3.2 Materials and method

Our focus is on phonological and phonetic responses to IP-final time pressure that originate from a stepwise decrease in sonorant duration of four monosyllabic target words. Recording six varieties of Dutch (including Standard Dutch and West Frisian) and three nuclear pitch accents (falls, rises and fall-rises) allows us to explore whether the observed responses are dialect- and/or contour-specific. The addition of the complex fall-rise in this investigation is of particular interest, first because it allows us to investigate whether it is avoided in some dialects, as we have seen for varieties of English and German, and second because its two  $f_0$  movements allow for many combinations of phonetic responses.

#### 3.2.1 Varieties and subjects

We made recordings in five places along the Dutch coast, covering different dialectal subgroups (Figure 3). Zeelandic Dutch in Zuid-Beveland (ZB), Southern Hollandic in Rotterdam (RO), and Northern Hollandic in Amsterdam (AM) all belong to the Low Franconian language family. We also recorded a West Frisian variety in Grou (GR) and a Low Saxon variety in Winschoten (WI). The Standard Dutch (SD) speakers were recorded in Nijmegen. Standard Dutch is most closely related to western varieties such as Rotterdam and Amsterdam (cf. Smakman 2006 and references therein).



Figure 3. Recording locations in the Netherlands.

A total of 119 speakers participated in the experiments, 49 of whom were male<sup>4</sup> (see table 1). They were aged between 14 and 49. Participants were university students (SD), secondary school students (ZB), members of a Scouting club (RO, AM) or members of the local community (GR, WI). The speakers from Zuid-Beveland, Grou, and Winschoten were bilingual with Standard Dutch and their local language. All regional speakers and at least one of their parents were raised in the selected place and spoke the indigenous variety fluently. For Standard Dutch, the procedure was different, as the area where this variety is spoken is less determined by geographical boundaries. Speakers could participate if they reported to speak Standard Dutch. Besides self-reporting, two Dutch phoneticians independently judged each recording<sup>5</sup>. Recordings were included if the judges agreed that the geographical and linguistic origin of the participants could not be determined by their accent. In other words, we did not allow for regional features to be present in our speakers' pronunciation<sup>6</sup>. Except for the speakers of West Frisian and Standard Dutch, our speakers were less familiar with their local language as a written language, which may have had a negative influence on the fluency of the speech in the reading task of some speakers.

---

<sup>4</sup> The distribution over gender is skewed particularly for those varieties where we recruited from the local town population (i.e., Winschoten and Grou), where it proved difficult to find male speakers.

<sup>5</sup> The combination of internal and external judgement seems a valid method if we consider the results of a perceptual distances study by Charlotte Gooskens (1997). In her study, six groups of listeners (one Standard Dutch group and five non-standard groups) judged the 'degree of Standard Dutch' of fragments of SD speech<sup>5</sup> on a scale from 1 to 10 (10 being 'ideal' SD). No statistical differences were found between the judgements of the six groups, meaning that both the standard and non-standard listeners agreed on the degree of SD of the speech fragments (mean scale position: 8.2). Gooskens (1997:76) concludes from these results that listeners, regardless their linguistic or geographical background, are quite able to recognize Standard Dutch speech, probably because they are familiar with it from the media.

<sup>6</sup> By doing so, our definition of Standard Dutch tends towards Smakman's (2006) 'exclusive' interpretation rather than the 'inclusive' interpretation, in which Standard Dutch is seen as the language that is used by a large group of first or second language speakers as a means of communication in which dialect features are excluded. The interpretation allows for some variation, which will be mostly in the pronunciation, as the language is almost completely standardised where other types of variation are concerned. Both regional and non-regional features will be present, as the main goal of the speakers is to be mutually understandable. In the exclusive interpretation, Standard Dutch is defined as a strictly homogenous (and somewhat unnatural) language. Its form is subject to a set of rules, and variation is allowed only within strict boundaries. According to the exclusive definition, the majority of the Dutch population are speakers of nonstandard varieties and only a small elite group speaks Standard Dutch.

Table 1. *Number, gender and average age (range) of speakers from Standard Dutch (SD), Zuid-Beveland (ZB), Rotterdam (RO), Amsterdam (AM), Grou (GR) and Winschoten (WI).*

	<i>female</i>	<i>male</i>	<i>total</i>	<i>average age</i>	<i>age range</i>	<i>average age female</i>	<i>average age male</i>
SD	14	9	<b>23</b>	<b>23</b>	<b>18-30</b>	23	23
ZB	8	10	<b>18</b>	<b>19</b>	<b>16-33</b>	21	17
RO	7	12	<b>19</b>	<b>22</b>	<b>16-31</b>	20	23
AM	7	11	<b>18</b>	<b>31</b>	<b>14-45</b>	33	30
GR	20	3	<b>23</b>	<b>28</b>	<b>17-43</b>	28	35
WI	14	4	<b>18</b>	<b>32</b>	<b>15-49</b>	30	41

Participants' recordings were excluded if they were (highly) disfluent or appeared to the experimenter not to speak naturally, if the speakers afterwards reported that they were dyslectic or had hearing problems, or if the speakers turned out not to satisfy the requirements with respect to their linguistic and/or geographical background. All participants were naive as to the purpose of the task and were paid for their participation.

### 3.2.2 Materials

We designed a sentence-reading task in which four fictitious monosyllabic proper names (*lof*, *loof*, *lom*, and *loom*) were used as target words in IP-final position. Their phonetic space for voicing was reduced in four steps by varying the phonological vowel length of the nucleus (V /ɔ/ vs. VV /o:/) and the sonorant nature of the coda (sonorant /m/ vs. voiceless /f/). The onset consonant /l/ was kept constant to avoid undesirable segmental effects. The four words were kept constant in all language versions.

Based on average durations of Standard Dutch vowels and consonants as reported in Waals (1999), we expected all four target words to differ significantly in terms of sonorant duration, with *lof* < *loof* < *lom* < *loom*. This design will in principle allow us to investigate whether the observed accommodation effects are gradient or categorical.

The set of twelve target sentences consisted of four statements, four yes/no questions and four rhetorical questions. We designed the statements to elicit a falling nuclear pitch contour on the IP-final target word (H\*L L%); the yes/no questions to elicit a nuclear rise (L\*H H%), and the rhetorical questions – which are syntactically statements – to elicit a fall-rise (H\*L H%). The rhythmic structure and the number of syllables preceding the target word was kept constant within each condition. To avoid possible alignment and/or scaling effects due to accent clash (cf. Silverman and Pierrehumbert 1990), the syllable immediately preceding the target word was unstressed. Similarly, to reduce the chances of speakers producing prenuclear accents, which can trigger Downstep (Pierrehumbert 1980, van den Berg, Gussenhoven, and Rietveld 1992) and

possibly other allophonic effects on the nuclear contours, the sentences were kept short and new information was given only in the final noun phrase. All target sentences were preceded by a context sentence with which they formed a mini-dialogue. The three types of mini-dialogues are listed in table 2 for the target word *lom*. The complete set of experimental sentences for all varieties is listed in the Appendix.

Table 2. Dutch context sentences and experimental sentences used to elicit falls, rises and fall-rises, with English translations. The target sentences are printed in bold; the word carrying the nuclear pitch accent is capitalized.

Sentence type	Context sentence (A)	Target sentence (B)
Statement	Met wie gingen de kinderen naar de dierentuin? <i>With whom did the children go to the zoo?</i>	<b>Ze gingen met meester LOM.</b> <i>They went with mister Lom.</i>
Yes/no-question	Ik zag net je broer Koen met je buurvrouw langslopen. <i>I just saw your brother Koen walking past with your neighbour.</i>	<b>Liep-ie naast mevrouw de LOM?</b> <i>Was he walking with Mrs. de Lom?</i>
Rhetorical question	Pepijn de Heer komt straks ook naar het feest. <i>Pepijn de Heer is also coming to the party.</i>	<b>Hij heet toch Pepijn de LOM?</b> <i>But he's called Pepijn de Lom, isn't he?</i>

### 3.2.3 Procedure

To avoid listing effects, the 12 mini-dialogues were interspersed with 61 filler sentences (used for other experiments) and presented in a booklet, in randomized order which was reversed for half of the subjects per variety. Speakers were recorded in pairs to reduce any effects of the experimenter's presence and the nature of the task on their dialect level. One speaker read the context sentence (A); the other the target sentence (B). The participants switched roles at the end of the task, after they had repeated any mispronounced sentences.

We collected the Standard Dutch data in a pilot experiment. A slightly modified version of the pilot sentences formed the Dutch set of experimental sentences, which was used for Zuid-Beveland, Rotterdam and Amsterdam. Speakers from Zuid-Beveland translated the Dutch sentences to their variety as they went along. Frisian and Low Saxon have their own standardized spelling system and we therefore translated the test materials to the local variety for Grou and Winschoten. For all varieties, the rhythmic, lexical, and segmental context was kept comparable to the Standard Dutch materials as much as possible.

Recordings of the local varieties were made in a quiet room either in the homes of our speakers or in a public building. The Standard Dutch recordings were made in a sound-treated booth at Radboud University Nijmegen. We used a portable digital recorder (Tascam HD P2 for Standard Dutch and Zoom H4 for

all other varieties) with a 48kHz sampling rate, 16 bit resolution and stereo format. Participants all wore Shure WH30XLR or Sennheiser MKE 2 wired condenser headset microphones.

### 3.2.4 Data selection and analysis

All recorded target sentences were converted to monaural files and stored on computer disk as separate wav. files. From each speaker and condition, one utterance per target sentence was selected for further analysis. If a speaker produced the same sentence with more than one nuclear contour type (e.g., once with a fall, and once with a downstepped fall), we selected the type that occurred most often within its pragmatic condition for that speaker's variety. Besides non-fluent utterances, we also excluded utterances in which the nuclear accent appeared in non-final position. We transcribed the nuclear pitch accent of each utterance following Gussenhoven (2005).

In all varieties except Amsterdam, the fall (H\*L L%) is the predominant nuclear contour for statements, which is used between 67% (Rotterdam) and 96% of cases (Winschoten)<sup>7</sup>. In Amsterdam, about half of the falling contours are realized with a late peak and a falling movement that doesn't reach the bottom of the speaker's range (L\*HL 0%). We find the downstepped fall in all varieties except Winschoten (used in 6-19%). The fall-rise (H\*L H%) is the preferred choice for the rhetorical questions (82-97%). The most striking result is found in the yes/no-questions, which are mostly realized with H\*L L% in Zuid-Beveland. We will use the label 'IF' for these interrogative falls, and the label 'DF' for the declarative falls used in statements. The other varieties produce rising and falling-rising nuclear contours for questions. Note that the rising contours consist of low rises (L\*H H%), high rises (H\* H%) and low low rises (L\* H%). While IP-final monosyllabic low low rises could in general be easily recognized on the basis of their late alignment in the syllable, it was hard to distinguish between low rises and high rises in this position. We therefore decided to collapse the data of low rises and high rises and to exclude the low low rises from our analyses.

Acoustic and auditory analysis of the data was done with the help of the speech processing software package Praat (Boersma and Weenink 1992-2008). We inserted the tonal and segmental labels in table 3.

---

<sup>7</sup> We refer to Hanssen, Gussenhoven and Peters (submitted, also see chapter 6) for a detailed account of melodic preferences and contour distributions over pragmatic contexts.

Table 3. Overview of tonal and segmental labels.

Tonal labels	Fall (F)	Rise (R)	Fall-Rise (FR)
L1	<i>n/a</i>	f0 elbow at beginning of nuclear rise	<i>n/a</i>
H1	nuclear f0 peak	end of nuclear rise	nuclear f0 peak
L2	end of nuclear fall	<i>n/a</i>	elbow between two f0 maxima
H2	<i>n/a</i>	<i>n/a</i>	end of nuclear fall-rise
Segmental labels			
V	beginning of nucleus of accented word		
C	beginning of coda of accented word		
E	end of coda of accented word		

Segment boundaries were determined on the basis of visual inspection of the broadband spectrogram and the waveform, along with auditory information (Turk, Nakai, and Sugahara 2006). We placed all labels at negative-to-positive zero-crossings. H1, H2 (falls, rises and fall-rises) and L2 (fall-rises) were placed semi-automatically using a *Praat* function that located the f0 maximum or minimum in a selected interval. Labelling L1 in rises, and L2 in falls semi-automatically<sup>8</sup> led to unreliable data, since the f0 minimum preceding or following the f0 maximum did not always correspond to the location of the elbow. These labels were therefore placed manually, by searching for the location with the fastest change in the speed of the f0 movement near the bottom line of the nuclear contour. Pitch tracking errors such as octave jumps were corrected by hand. See Del Giudice et al (2007) and Petrone and D’Imperio (2009) for an overview of manual and (semi-)automatic labelling procedures.

A *Praat Script* computed and saved the f0 value (f) and time (t) of each label. To neutralize gender differences in f0 excursion, we converted the frequency values from Hz to semitones (ST re 100 Hz) and computed the dependent variables in Table 4.

We randomly selected 144 sentences (12 items x 2 speakers x 6 varieties) to be labelled independently by two trained phoneticians. The results of a reliability test to check the inter-labeler agreement are presented in table 5, which shows the mean differences between the two measurements for each of the labels listed in table 3. Inter-labeler agreement was at least 0.998 (Cronbach’s alpha) for all labels.

<sup>8</sup> Neither the semi-automatic *Praat* function, nor the Elbow Script (Beckman and Welby 2006), yielded reliable results.

Table 4. *An overview of the dependent variables calculated from the tonal and segmental labels. Part 1 means the falling part of fall-rise and part 2 the rising part.*

Variables	Description	Formula	F	FR	R
<i>Durational</i>					
SONRIME	Duration of sonorant time (ms)	$t(E) - t(V)$ for <i>lom, loom</i> $t(C) - t(V)$ for <i>lof, loof</i>	✓	✓	✓
<i>Timing</i>					
PEAKDELAY	Distance of nuclear peak to beginning of nuclear rime (ms)	$t(H1) - t(V)$	✓	✓	
RELPEAK	Location of nuclear peak as a proportion of sonorant rime duration (%)	$(PEAKDELAY / SONRIME) * 100$	✓	✓	
VALLEYDELAY	Distance of nuclear valley to beginning of nuclear rime (ms)	$t(L1) - t(V)$			✓
RELVALLEY	Location of nuclear valley as a proportion of sonorant rime duration (%)	$(VALLEYDELAY / SONRIME) * 100$			✓
<i>Shape</i>					
F0DUR	Duration of f0 movement (ms)	$t(L2) - t(H1)$ for <i>fall</i> $t(H1) - t(L1)$ for <i>rise</i> $t(H2) - t(H1)$ for <i>fall-rise</i>	✓	✓	✓
F0EXC	Excursion of f0 movement (ST)	$f0(L2) - f0(H1)$ for <i>fall</i> $f0(H1) - f0(L1)$ for <i>rise</i>	✓		✓
ROFCH	Rate of f0 change; the average speed of the f0 movement (ST/100ms)	$F0EXC / F0DUR * 100$	✓		✓
F0DUR_1, F0DUR_2	Duration of Part 1 and 2 of <i>fall-rise</i> (ms)	$t(L2) - t(H1)$ for Part 1, $t(H2) - t(L2)$ for Part 2		✓	
F0EXC_1, F0EXC_2	Excursion of Part 1 and 2 of <i>fall-rise</i> (ST)	$f0(L2) - f0(H1)$ for Part 1, $f0(H2) - f0(L2)$ for Part 2		✓	
ROFCH_1, ROFCH_2	Rate of f0 change of Part 1 and 2 of <i>fall-rise</i> (ST/100ms)	$F0EXC\_1 / F0DUR\_1 * 100$ for Part 1, $F0EXC\_2 / F0DUR\_2 * 100$ for Part 2		✓	

Speakers' data were included if they had used the intended nuclear contour for at least three of the four target words, i.e. H\*L L% for statements; L\*H H% or H\* H% for yes/no-questions, and H\*L H% for rhetorical questions. Table 6 gives the number of remaining speakers per analysis, and table 7 shows the proportion of missing data cells per word and contour condition.

Table 5. *Results of reliability test of the timing of 8 labels by 2 labelers.*

Label	Cronbach's Alpha	Mean difference in ms
V	1.000	-1.77
C	1.000	5.89
E	0.999	3.28
L1	0.998	-6.00
H1	0.998	0.68
L2	0.999	1.98
H2	1.000	1.21

Table 6. *Number of speakers recorded for Standard Dutch (SD), Zuid-Beveland (ZB), Rotterdam (RO), Amsterdam (AM), Gron (GR) and Winschoten (WI), and number of speakers used per sentence condition (statements, yes/no-questions and rhetorical questions).*

	No speakers recorded	Statements	Yes/no- questions	Rhetorical questions
SD	23	17	11	22
ZB	18	15	13 <sup>a)</sup>	8
RO	19	13	12	15
AM	18	6	9	12
GR	23	21	6	22
WI	18	16	11	15
total	119	88	62	94

<sup>a)</sup> These thirteen speakers from ZB used H\*L L% for at least three of the four test words in yes/no-questions, as opposed to a rising nuclear contour, and will therefore be analyzed in the category 'falls' from now on. Falls used for statements are labeled DF ('declarative fall') and falls used for yes/no-questions are labeled IF ('interrogative fall').

Table 7. *Proportion of missing cells per word and contour condition for each variety.*

	FALLS				RISES				FALL-RISES			
	lof	loof	lom	loom	lof	loof	lom	loom	lof	loof	lom	loom
SD	.18	.06	.06	.00	.00	.00	.09	.45	.27	.00	.00	.00
ZB_DF	.13	.07	.00	.13					.38	.00	.00	.00
ZB_IF	.00	.00	.23	.15								
RO	.15	.15	.00	.15	.00	.25	.17	.33	.47	.00	.00	.00
AM	.17	.17	.00	.00	.00	.11	.33	.00	.25	.00	.00	.00
GR	.10	.00	.05	.14	.00	.17	.00	.27	.14	.00	.00	.05
WI	.13	.00	.00	.06	.27	.00	.00	.27	.07	.00	.00	.00
total	.12	.05	.05	.10	.06	.10	.12	.27	.24	.00	.00	.01



### 3.2.5 Statistical and visual analysis

We analyzed the data using the Linear Mixed Effect Model in SPSS, including SPEAKER as random factor, and WORD (*lof*, *loof*, *lom*, *loom*) as fixed factor. Pairwise comparisons between the four levels of the fixed factor were carried out using the Bonferroni correction.

To estimate the additional amount of variance explained by adding the fixed factor WORD to the model, as opposed to a model that only includes the random factor SPEAKER, we used  $\Omega^2$ , following Xu (2003). The formula is

$$\Omega^2 = 1 - \frac{\text{variance residuals model random \& fixed}}{\text{variance residuals model random}}$$

The statistical analysis is based on the time and f0 value of tonal and segmental targets. To be able to visually display general variation in contour shape between tonal targets, we generated images of averaged f0 contours, using a *Praat Script* that measures f0 values (in ST re 100 Hz) of 100 data points in a selected interval. We produced averaged f0 contours separately for each contour condition, dialect and target word.

## 3.3 Results

In section 3.3.1, we check whether our experimental manipulation worked, and whether the sonorant portion of our experimental items indeed decreased in four steps (*loom* > *lom* > *loof* > *lof*). We then present our results, in four steps that follow the possible responses to time pressure outlined in the introduction. In section 3.3.2, we address the question whether speakers avoid complex contours in durationally unfavorable situations. Section 3.3.3 reports whether speakers take more time to execute more complex contours in terms of segmental duration. Section 3.3.4 reports the effects of increasing time pressure on the alignment of tonal targets. Finally, section 3.3.5 reports the effects of time pressure on the shape of nuclear contours, i.e. on their f0 duration, excursion and slope, and on the scaling of individual tonal targets.

### 3.3.1 Sonorant rime duration

We measured the duration of the sonorant rime of the four target words to check whether they show a stepwise increase in available sonorant material. As table 8 shows, we found a main effect of the fixed factor WORD on SONRIME in all varieties and all contour conditions. The bar charts in Figure 4 illustrate that overall, a stepwise increase in sonorant rime duration from *lof* to *loom* is clearly present for all contours. However, the results of pairwise comparisons between adjacent word pairs show that *loof* and *lom* are often not significantly different, particularly in rises. We conclude that our manipulation was generally successful,

but that we expect time pressure effects to surface mostly between word pairs 1 and 2/3, and between word pairs 2/3 and 4, and less so between 2 and 3.

Table 8. *Effects and pairwise comparisons (Bonferroni) of WORD on sonorant rime duration (SONRIME) in falls, rises and fall-rises.*

SONRIME		$p$	$\Omega^2$	1-2	1-3	1-4	2-3	2-4	3-4
FALLS									
SD	$F(3,44)=68.67$	***	.843	***	***	***	*	***	***
ZB_DF	$F(3,37)=85.95$	***	.881	***	***	***	*	***	***
ZB_IF	$F(3,32)=75.26$	***	.872	***	***	***	ns	***	***
RO	$F(3,30)=156.99$	***	.950	***	***	***	***	***	***
AM	$F(3,13)=61.06$	***	.946	**	***	***	ns	***	**
GR	$F(3,54)=116.88$	***	.889	***	***	***	ns	***	***
WI	$F(3,42)=95.29$	***	.890	***	***	***	**	***	***
SONRIME		$p$	$\Omega^2$	1-2	1-3	1-4	2-3	2-4	3-4
FALL- RISES									
SD	$F(3,58)=98.84$	***	.869	***	***	***	***	***	***
ZB	$F(3,18)=19.61$	***	.772	ns	***	***	*	**	ns
RO	$F(3,36)=74.38$	***	.888	***	***	***	***	***	***
AM	$F(3,30)=71.85$	***	.905	***	***	***	*	***	***
GR	$F(3,59)=128.38$	***	.892	***	***	***	ns	***	***
WI	$F(3,41)=51.11$	***	.805	**	***	***	ns	***	***
SONRIME		$p$	$\Omega^2$	1-2	1-3	1-4	2-3	2-4	3-4
RISES									
SD	$F(3,25)=48.62$	***	.863	***	***	***	ns	***	***
RO	$F(3,25)=73.09$	***	.912	***	***	***	ns	***	***
AM	$F(3,21)=60.01$	***	.901	***	***	***	ns	***	***
GR	$F(3,13)=21.62$	***	.856	***	*	***	ns	ns	*
WI	$F(3,25)=22.44$	***	.746	***	***	***	ns	**	*

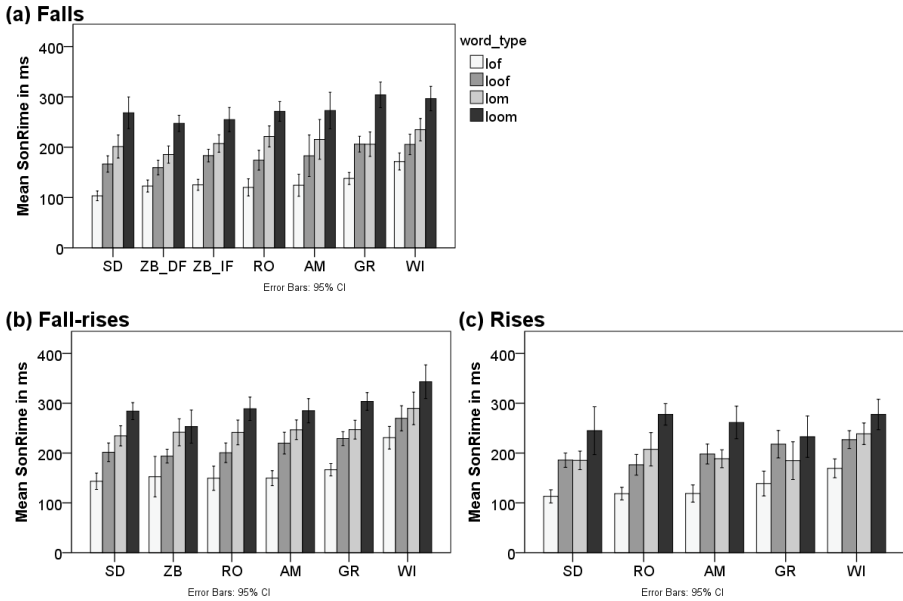


Figure 4. Sonorant rime duration in ms, for the four target words. Panel a represents the fall and includes the averages of the Zuid-Beveland interrogative falls (ZB\_IF). Panels b and c represent the rises and fall-rises, respectively. Error bars represent the 95% confidence interval.

### 3.3.2 Avoidance of complex contours

The common use of the fall-rise for rhetorical questions reported in section 3.2.4 shows that complex contours are not structurally avoided in IP-final position in any of the varieties. Apparently, the adverb *toch* is such a strong trigger of the falling-rising melody (cf. van de Ven and Gussenhoven 2011) that speakers consistently produce a complex contour instead of a simple one, even with limited sonorant material at their disposal. Yet, as table 9 illustrates, we still find that their frequency is affected by the proximity of the final boundary. The proportion of simple contours in rhetorical questions in Standard Dutch, Rotterdam, and Amsterdam is larger for the shortest word condition *lof* than for longer words. The proportion of simple contours in Zuid-Beveland, Grou and Winschoten is low, regardless of target word. The proportion of simple contours for yes/no-questions is larger overall, and differentiated by target word duration most clearly in Standard Dutch, Zuid-Beveland and Amsterdam, with shorter words attracting higher proportions.

Table 9 *Proportion of simple contours for each target word in rhetorical questions (RQ) and yes/no-questions (YN).*

	RQ						YN					
	SD	ZB	RO	AM	GR	WI	SD	ZB	RO	AM	GR	WI
<i>lof</i>	.18	.06	.21	.25	.05	.06	.86	1.00	.95	.94	.68	.72
<i>loof</i>	.00	.06	.06	.00	.00	.00	.73	1.00	.89	.83	.78	.89
<i>lom</i>	.04	.06	.06	.13	.00	.11	.74	.89	.83	.59	.61	.83
<i>loom</i>	.00	.06	.05	.12	.05	.00	.48	.87	.94	.64	.77	.61

### 3.3.3 Durational differences between complex vs simple contours

Compared to the fall (H\*L L%), the fall-rise (H\*L H%) is more complex since it involves the production of three instead of two pitch movements. A possible response to this type of time pressure is to increase the sonorant duration of the syllable on which the contour is produced. Table 10 shows that CONTOUR (simple vs complex) had a main effect on SONRIME in all varieties<sup>9</sup>. Sonorant rime duration in fall-rises was always significantly longer than in falls. This is also illustrated in the bar charts in Figure 4 (panel a and b).

Table 10. *Effects of CONTOUR (F vs FR) on sonorant rime duration (SONRIME).*

SONRIME		<i>p</i>	$\Omega^2$	F mean	FR mean
SD	$F(1,122) = 22.18$	***	.157	191	215
ZB <sup>a)</sup>	$F(1,73) = 23.25$	***	.347	179	207
RO	$F(1,81) = 7.00$	**	.082	199	213
AM	$F(1,54) = 13.65$	***	.250	201	227
GR <sup>b)</sup>	$F(1,135) = 27.86$	***	.244	213	235
WI	$F(1,98) = 97.70$	***	.497	229	278

<sup>a)</sup> The WORD x CONTOUR interaction  $F(3,61) = 4.40$   $p < 0.01$  in Zuid-Beveland may be explained by the fact that SONRIME in *loom* differed by only 4 ms between falls (247 ms) and fall-rises (251 ms). This difference is much bigger for the other target words.

<sup>b)</sup> There are two possible explanations of the WORD x CONTOUR interaction  $F(3,132) = 4.58$   $p < 0.01$  in Grou. First, SONRIME varies between *loof* and *lom* in fall-rises but not in falls. Secondly, there's no difference between SONRIME of *loom* in falls and fall-rises (305 vs 304 ms), whereas SONRIME for the other target words is always considerably longer in fall-rises than in falls.

### 3.3.4 Timing adjustments under time pressure

Under time pressure due to contour complexity, speakers may start the contour earlier in fall-rises compared to falls, to create more time for its execution. This is illustrated in section 3.3.4.1. Similarly, the timing of tonal targets may generally be affected by time pressure due to the proximity of an upcoming IP-boundary.

---

<sup>9</sup> As we are interested in the difference between falls and fall-rises, we report the results for the factor CONTOUR, and mention the results for the factor WORD only when there is a meaningful interaction between CONTOUR and WORD.

We report these effects separately for falls, fall-rises and rises in sections 3.3.4.2 – 3.3.4.4.

### 3.3.4.1 Simple vs complex contours

Table 11 shows a main effect of CONTOUR on RELPEAK in Standard Dutch, Grou and Winschoten, with the peak starting earlier in the sonorant rime in fall-rises. The bar charts in Figure 5 illustrate this result and show that differences in Rotterdam and particularly in Zuid-Beveland are small or non-existent. We will return to this result in the discussion.

Table 11. *Effects of CONTOUR (F vs FR) on relative peak alignment (RELPEAK).*

RELPEAK		<i>p</i>	$\Omega^2$	F mean	FR mean
SD	$F(1,126) = 11.92$	***	.128	25.69	18.33
ZB	$F(1,71) = 1.53$	ns	.045	17.39	20.65
RO	$F(1,86) = 0.29$	ns	.110	25.38	24.43
AM	$F(1,59) = 1.47$	ns	.097	20.85	17.66
GR <sup>a)</sup>	$F(1,137) = 40.81$	***	.330	28.30	19.06
WI	$F(1,98) = 4.88$	*	.068	26.08	22.25

<sup>a)</sup> The WORD  $\times$  CONDITION interaction  $F(3,132) = 3.03$   $p < 0.05$  in Grou can possibly be explained by the fact that the differences in RELPEAK between falls and fall-rises are larger for target words *lof* and *lom* than for *loof* and *loom*. Alternatively, the interaction may be caused by the fact that *loof* has a much later peak than the other target words in fall-rises, but not in the falls.

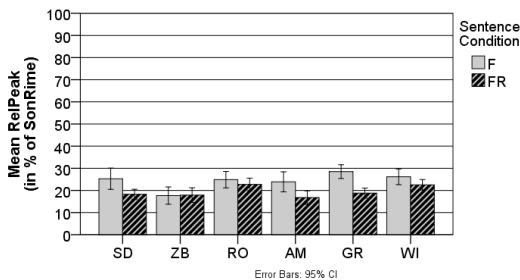


Figure 5. Relative peak alignment (in % of sonorant rime duration) in falls and fall-rises. Error bars represent the 95% confidence interval.

### 3.3.4.2 Falls

Table 12 shows that WORD had a significant effect on the timing of the peak (PEAKDELAY), measured relative to the beginning of the nucleus, in all varieties. The results of pairwise comparisons, along with inspection of the bar charts in Figure 6, suggest that the main effect is caused by significantly earlier peaks in *lof*

compared to the longer words in SD, ZB\_IF, RO and GR, and significantly later peaks in *loom* compared to the shorter words in ZB\_DF and WI.

Panel (a) of Figure 6 shows that there are large differences in absolute peak alignment between the four target words. If peak timing is measured as a proportion of the sonorant rime, as in panel (b), the differences are much smaller. Indeed, as table 12 shows, relative peak location (RELPEAK) is not affected by WORD in any of the varieties except Rotterdam. Speakers time the location of the peak with respect to the duration of the sonorant rime, and place it around 20 to 30% into the syllable rime. In Rotterdam, the peak location is unexpectedly late in *loof* and *lof*.

Table 12. *Effects and pairwise comparisons of WORD on absolute (PEAKDELAY) and relative peak location (RELPEAK) in falls.*

PEAK DELAY		<i>p</i>	$\Omega^2$	1-2	1-3	1-4	2-3	2-4	3-4
SD	$F(3,43) = 10.31$	***	.431	**	***	***	ns	ns	ns
ZB_DF	$F(3,38) = 8.17$	***	.397	ns	ns	***	ns	*	**
ZB_IF	$F(3,32) = 9.67$	***	.511	**	**	***	ns	ns	ns
RO	$F(3,31) = 5.03$	**	.340	*	ns	**	ns	ns	ns
AM	$F(3,13) = 6.20$	**	.586	*	*	*	ns	ns	ns
GR	$F(3,55) = 10.242$	***	.399	**	**	***	ns	ns	ns
WI	$F(3,42) = 5.23$	**	.279	ns	ns	**	ns	ns	*
REL PEAK		<i>p</i>	$\Omega^2$	1-2	1-3	1-4	2-3	2-4	3-4
SD	$F(3,42) = 1.98$	ns	.122	ns	ns	ns	ns	ns	ns
ZB_DF	$F(3,38) = 0.86$	ns	.047	ns	ns	ns	ns	ns	ns
ZB_IF	$F(3,32) = 1.05$	ns	.082	ns	ns	ns	ns	ns	ns
RO	$F(3,31) = 5.67$	**	.390	ns	ns	ns	*	*	ns
AM	$F(3,14) = 2.60$	ns	.312	ns	ns	ns	ns	ns	ns
GR	$F(3,55) = 2.15$	ns	.107	ns	ns	ns	ns	ns	ns
WI	$F(3,42) = 1.47$	ns	.086	ns	ns	ns	ns	ns	ns

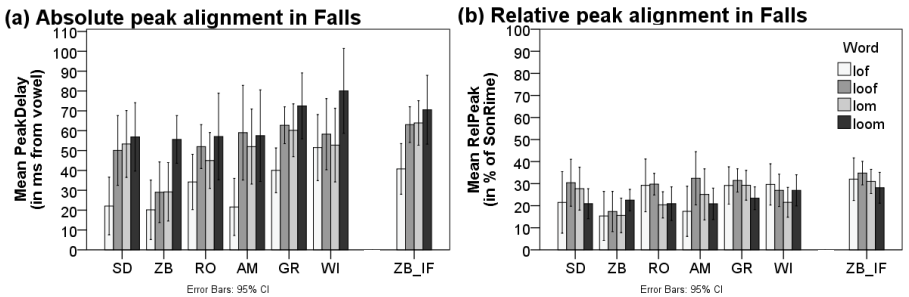


Figure 6. Absolute peak alignment (in ms from nucleus, panel a) and relative peak alignment (in % of sonorant rime, panel b), in falls. Error bars represent the 95% confidence interval.

### 3.3.4.3 Fall-rises

As table 13 shows, WORD had a main effect on the absolute timing of the peak in all varieties. The bar charts in panel (a) of Figure 7 reveal a tendency for peaks in shorter words to occur earlier, yet this is not a regular response in all varieties. The bar charts, along with posthoc comparisons between the four words show that in Standard Dutch, the timing of the peak is earlier on shorter target words, with all word pairs significantly different except *loof* and *lom*. The same general pattern is found in ZB, with the peak in *loom* significantly later than in *lof* and *loof*, and in RO, where the peak in *loom* occurs significantly later than the other target words. In AM and GR, the peak aligned earlier in short-vowel words *lof* and *lom* than in long-vowel words *loof* and *loom*. Pairwise comparisons in GR confirm that the short-vowel words are always significantly different from the long-vowel words, but not from each other. In Winschoten, finally, peak timing is particularly late in *loom* compared to the shorter words, as in Rotterdam.

To see whether peak timing is related to rime duration, as in the falls, we also measured the location of the peak as a proportion of the sonorant rime. As table 13 shows, we found a main effect of WORD on RELPEAK in Standard Dutch, Rotterdam and Grou. Posthoc comparisons show that in each of these varieties, the main effect of WORD is caused by one significantly different word or word pair. In Standard Dutch, relative peak location in *lof* is significantly earlier than in *loom*; in Rotterdam, *lof* has a significantly later relative peak location than *loom*; and in Grou, the peak in *loof* is aligned later than in all other target words.

Table 13. *Effects and pairwise comparisons for WORD on absolute (PEAKDELAY) and relative peak location (RELPEAK) in fall-rises.*

PEAKDELAY		<i>p</i>	$\Omega^2$	1-2	1-3	1-4	2-3	2-4	3-4
SD	$F(3,58) = 24.69$	***	.575	**	***	***	ns	***	**
ZB	$F(3,18) = 5.92$	**	.514	ns	ns	*	ns	*	ns
RO	$F(3,35) = 5.44$	**	.325	ns	ns	*	ns	*	*
AM	$F(3,30) = 4.34$	*	.307	ns	ns	*	ns	ns	ns
GR	$F(3,59) = 19.21$	***	.508	***	ns	***	***	ns	***
WI	$F(3,41) = 9.22$	***	.419	ns	ns	**	ns	***	**
RELPEAK		<i>p</i>	$\Omega^2$	1-2	1-3	1-4	2-3	2-4	3-4
SD	$F(3,58) = 2.98$	*	.124	ns	ns	*	ns	ns	ns
ZB	$F(3,19) = 1.44$	ns	.187	ns	ns	ns	ns	ns	ns
RO	$F(3,35) = 4.19$	*	.274	ns	*	*	ns	ns	ns
AM	$F(3,30) = 2.13$	ns	.176	ns	ns	ns	ns	ns	ns
GR	$F(3,59) = 11.06$	***	.375	**	ns	ns	***	*	ns
WI	$F(3,41) = 2.28$	ns	.143	ns	ns	ns	ns	ns	ns

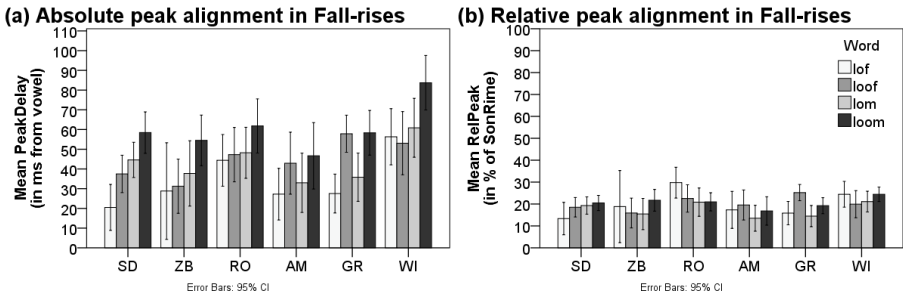


Figure 7. Absolute peak alignment (in ms from nucleus, panel a) and relative peak alignment (in % of sonorant time duration, panel b), in falls-rises. Error bars represent the 95% confidence interval.

### 3.3.4.4 Rises

As table 14 illustrates, we found a main effect of WORD on the start of absolute valley timing in all varieties. Unlike in the falls and fall-rises, we also found a main effect of WORD on proportional valley timing in all varieties except Winschoten. For this variety, the bar charts in Figure 8 and posthoc comparisons show that the rise starts significantly earlier in *lof* than in all other words. Posthoc comparisons (Bonferroni) show that in all other varieties, neither *lof* and *loom*, nor *loof* and *loom* are significantly different, so that the effect may possibly be attributed to the difference between these pairs. The bar charts in Figure 8 show that the beginning of the rises in *lof* and *loom* are aligned before or shortly after the beginning of the vowel, while those in *loof* and *loom* are aligned later. These observations are true regardless of whether peak alignment is measured in absolute or relative terms. This issue will be further explored in the discussion (section 3.4.2).

Table 14. *Effects and pairwise comparisons for WORD on VALLEYDELAY and relative valley location (RELVALLEY) in rises.*

VALLEYD ELAY		<i>p</i>	$\Omega^2$	1-2	1-3	1-4	2-3	2-4	3-4
SD	$F(3,24) = 8.22$	***	.580	***	ns	*	*	ns	ns
RO	$F(3,25) = 4.76$	**	.367	ns	ns	ns	ns	ns	ns
AM	$F(3,28) = 8.17$	***	.467	*	ns	**	*	ns	**
GR	$F(3,13) = 20.60$	***	.842	**	ns	***	**	ns	***
WI	$F(3,26) = 5.14$	**	.381	**	*	*	ns	ns	ns
REL VALLEY		<i>p</i>	$\Omega^2$	1-2	1-3	1-4	2-3	2-4	3-4
SD	$F(3,24) = 4.92$	**	.442	*	ns	ns	ns	ns	ns
RO	$F(3,25) = 3.57$	*	.295	ns	ns	ns	ns	ns	ns
AM	$F(3,22) = 7.83$	***	.491	**	ns	*	*	ns	ns
GR	$F(3,13) = 13.45$	***	.775	*	ns	***	ns	ns	**
WI	$F(3,26) = 2.94$	ns	.260	*	ns	ns	ns	ns	ns



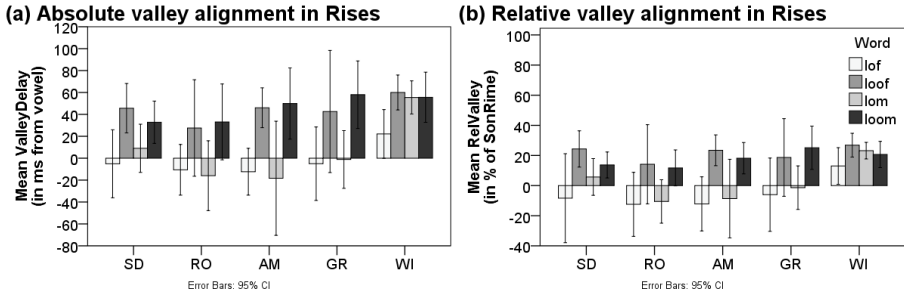


Figure 8. Valley alignment in rises ( $L^*H$   $H\%$  and  $H^*$   $H\%$ ). Panel (a) shows peak alignment relative to the beginning of the vowel in ms (VALLEYDELAY); panel (b) shows peak alignment as a proportion of the sonorant time in % (RELVALLEY). The beginning of the vowel is set to 0 ms and 0%, respectively. Error bars represent 95% confidence interval.

### 3.3.5 Shape adjustments under time pressure

In the final results section, we report how speakers adjust the shape of the nuclear pitch contour in terms of truncation or compression. We present the results separately for falls (section 3.3.5.1), fall-rises (3.3.5.2) and rises (3.3.5.3). For each contour condition, we look at the duration of the movement ( $f_0$  duration), its  $f_0$  excursion and slope. Since differences in excursion do not say anything about the height of individual pitch targets, we also look at tonal scaling. An overview of the results is shown in Figures 9-11.

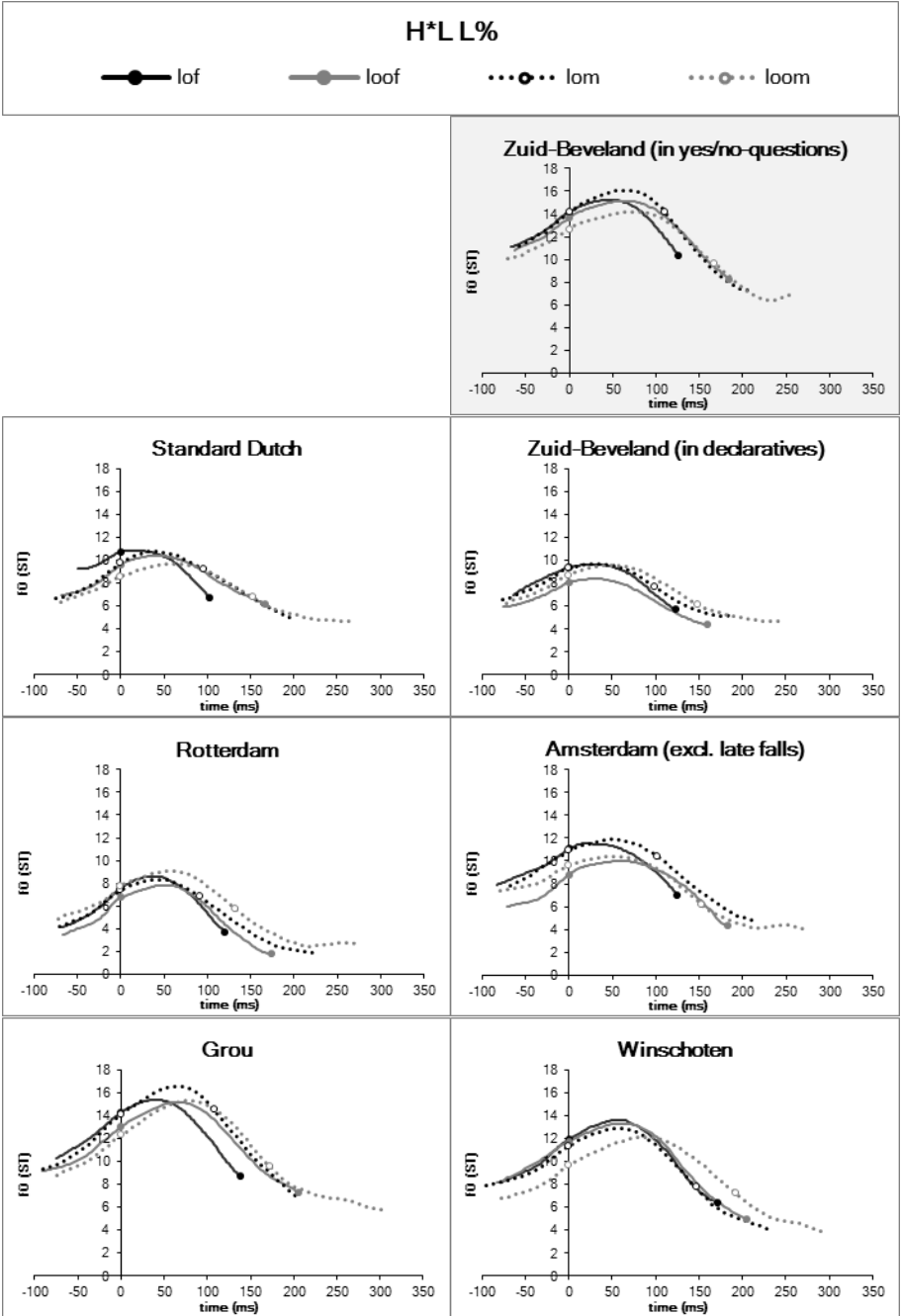


Figure 9. Averaged contours of four target words in all varieties for H\*L L% in declaratives and for ZB in yes/no-questions (gray panel). The beginning of the nuclear vowel is set to 0 ms. Circles represent the beginning and end of the nuclear vowel (/ə/ or /o:/).

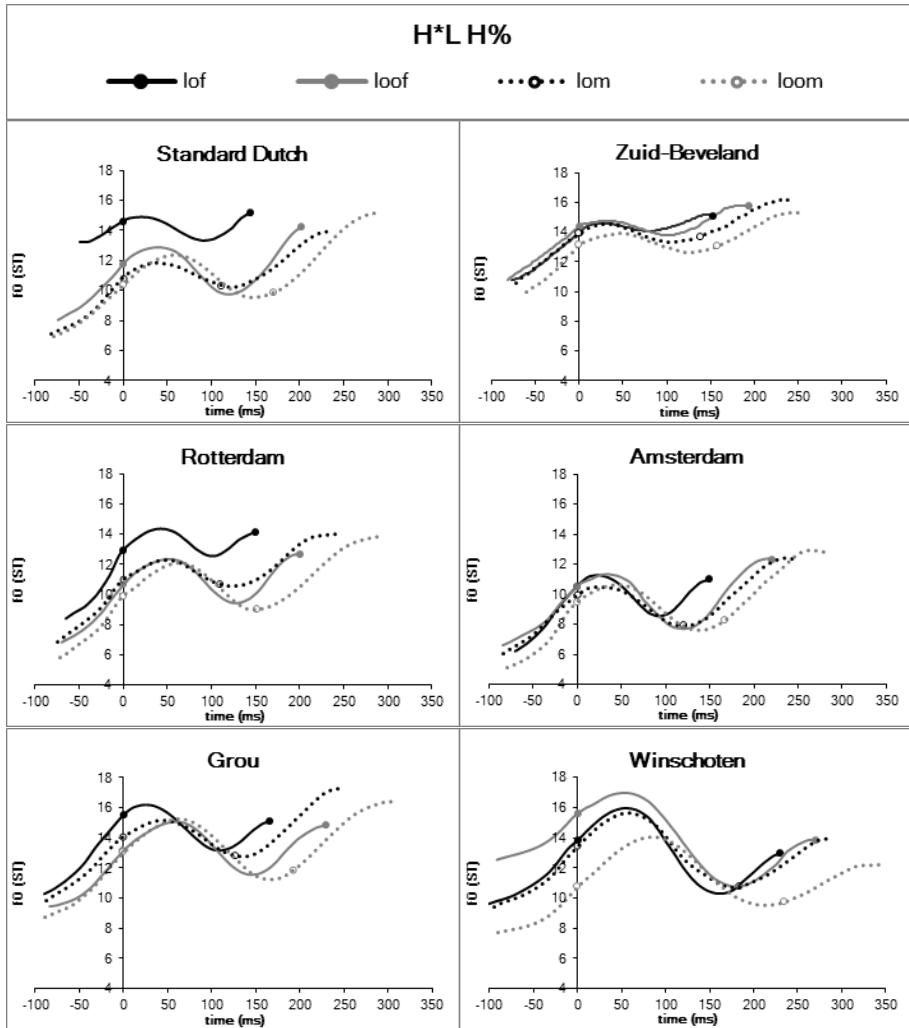


Figure 10. Averaged contours of four target words in all varieties for H\*L H% in rhetorical questions. The beginning of the nuclear vowel is set to 0 ms. Circles represent the beginning and end of the nuclear vowel (/ɔ/ or /o:/).

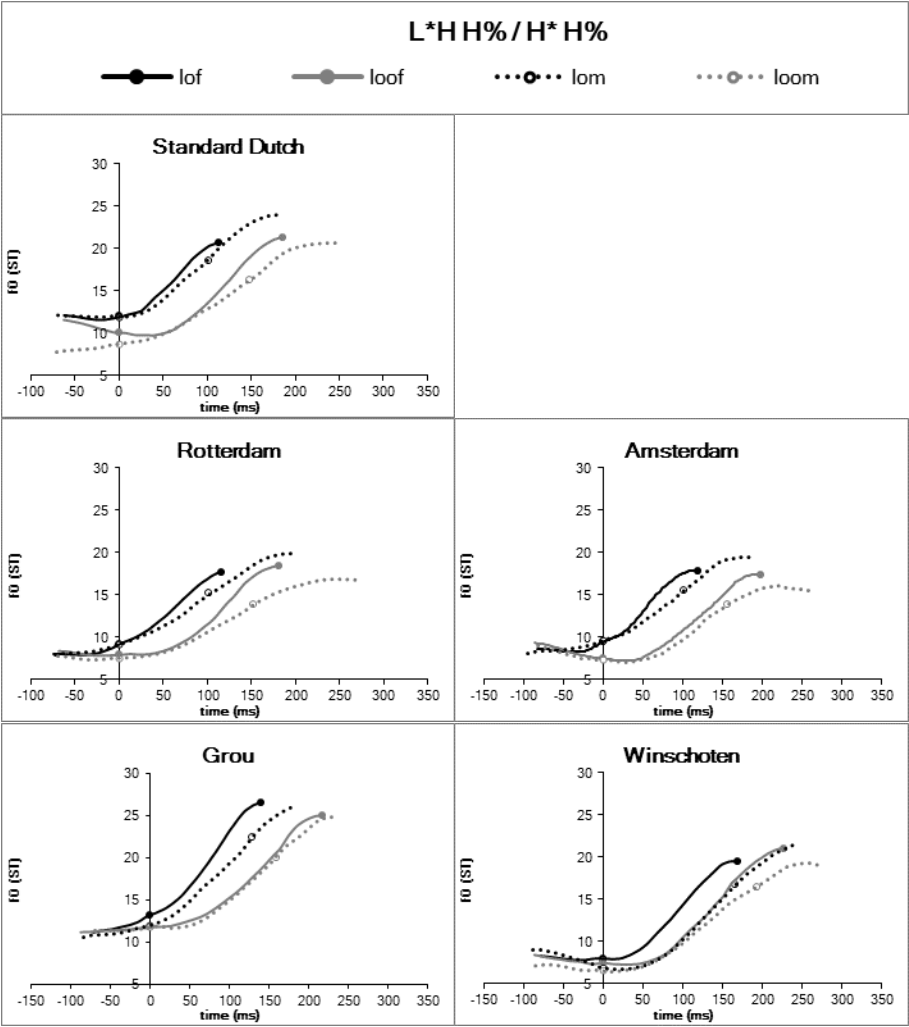


Figure 11. Averaged contours of four target words in all varieties except ZB for L\*H H% / H\* H% in yes/no-questions. The beginning of the nuclear vowel is set to 0 ms. Circles represent the beginning and end of the nuclear vowel (/ɔ/ or /o:/).

### 3.3.5.1 Falls

A central issue in contour adjustments is the extent to which increases in  $f_0$  duration match increases in sonorant rime duration. As table 15 shows, we found a main effect of WORD on  $f_0$  duration in all varieties except Amsterdam. Pairwise comparisons show that the largest differences occur between words with the shortest sonorant rime (*lof*) and longer sonorant rimes (*lom*, *loom*), as shown by the bar charts for  $f_0$  duration in Figure 13. However, a comparison of this figure to Figure 4 for sonorant rime duration shows that variation between the four target words is considerably larger for sonorant rime duration. Averaged across varieties, the difference between *lof* and *loom* is 38 ms for  $f_0$  duration, compared to 145 ms for sonorant rime duration. The proportion of the sonorant rime that is used to realize the falling movement from  $H^*$  to L is 47% for *loom*, 59% for both *lom* and *loof*, and 70% for *lof*. This is illustrated in Figure 12, which shows  $f_0$  duration within the sonorant syllable rime per word, averaged over all varieties. The part of the rime that is used to rise to the nuclear peak, and especially to level out after the end of the fall, also decreases as the rime gets shorter.

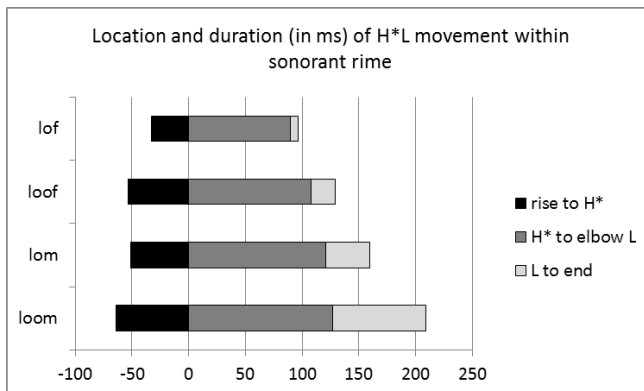


Figure 12.  $f_0$  duration from nuclear peak set off against sonorant rime duration in ms per word, averaged across all varieties. The black bars represent duration from vowel to nuclear peak, the dark gray bars represent the duration of the nuclear fall, and the light gray bars represent the elbow to the end of the sonorant rime, or the part of the nuclear contour where  $f_0$  levels out. All three bars together represent sonorant rime duration. The beginning of the nuclear fall is set to 0 ms.

Since the differences in  $f_0$  duration among target words are small compared to the differences in sonorant rime duration, the differences in the excursion and slope of the contours can also be expected to be small. The results of the Linear Mixed Effects model for fixed factor F0EXC in table 15 show that WORD has a significant effect on  $f_0$  excursion in Standard Dutch, Grou and Zuid-Beveland interrogative falls, with pairwise comparisons showing significant differences mainly between the shortest word *lof* and longer words *lom* and/or *loom*. For rate

of change (ROFCH in table 15), we found an effect of WORD only in Winschoten, with *lof* having a higher rate of change than *loom*.

The bar charts in Figure 13 show mean f0 duration, f0 excursion and rate of change, broken down by word and variety. Although patterns show irregularities across varieties, we can see that Winschoten stands out because of its stable f0 excursion and higher rates of change on shorter words. In all other varieties, shorter words tend to have shorter f0 durations, smaller f0 excursions and therefore fairly stable rates of change. Recall from Figure 1 that a stable rate of change indicates truncation of the fall or undershoot of the high target, and a higher rate of change is indicative of compression. As such, the results suggest that in response to time pressure, speakers of Winschoten are inclined to compress falls, whereas the other varieties are inclined to truncate the fall or undershoot the nuclear peak.

The results in table 15 for H\_SCALING (main effect of WORD in one variety) and L2\_SCALING (main effect of WORD in four varieties) point in the direction of truncation, more so than undershoot. This conclusion is strengthened by looking at the bar charts in Figure 13, which shows a tendency for shorter words to have a higher scaling of the elbow, and no evidence for lower scaling of the peak.

Table 15. *Effects and pairwise comparisons for WORD on f0 duration (F0DUR), f0 excursion (F0EXC), rate of f0 change (ROFCH), height of the peak (H\_Scaling) and of the subsequent elbow (L2\_Scaling) in falls.*

F0DUR		<i>p</i>	$\Omega^2$	1-2	1-3	1-4	2-3	2-4	3-4
SD	$F(3,44) = 13.36$	***	.519	ns	***	***	**	**	ns
ZB_DF	$F(3,38) = 4.03$	*	.286	ns	*	ns	ns	ns	ns
ZB_IF	$F(3,32) = 13.24$	***	.567	**	**	***	ns	ns	ns
RO	$F(3,31) = 7.88$	***	.492	ns	***	**	ns	ns	ns
AM	$F(3,13) = 2.85$	ns	.400	ns	ns	ns	ns	ns	ns
GR	$F(3,56) = 6.48$	**	.291	*	**	**	ns	ns	ns
WI	$F(3,42) = 6.50$	**	.298	ns	ns	***	ns	ns	ns
F0EXC		<i>p</i>	$\Omega^2$	1-2	1-3	1-4	2-3	2-4	3-4
SD	$F(3,44) = 3.98$	*	.232	ns	*	ns	ns	ns	ns
ZB_DF	$F(3,37) = 2.75$	ns	.176	ns	ns	ns	ns	ns	ns
ZB_IF	$F(3,31) = 5.59$	**	.374	ns	*	**	ns	ns	ns
RO	$F(3,31) = 2.82$	ns	.224	ns	ns	ns	ns	ns	ns
AM	$F(3,13) = 1.16$	ns	.199	ns	ns	ns	ns	ns	ns
GR	$F(3,55) = 4.21$	**	.191	ns	**	ns	ns	ns	ns
WI	$F(3,42) = .16$	ns	.002	ns	ns	ns	ns	ns	ns
ROFCH		<i>p</i>	$\Omega^2$	1-2	1-3	1-4	2-3	2-4	3-4
SD	$F(3,44) = 1.31$	ns	.081	ns	ns	ns	ns	ns	ns
ZB_DF	$F(3,37) = 1.03$	ns	.072	ns	ns	ns	ns	ns	ns
ZB_IF	$F(3,32) = .18$	ns	.022	ns	ns	ns	ns	ns	ns
RO	$F(3,31) = .64$	ns	.054	ns	ns	ns	ns	ns	ns
AM	$F(3,13) = .77$	ns	.147	ns	ns	ns	ns	ns	ns
GR	$F(3,55) = 1.84$	ns	.086	ns	ns	ns	ns	ns	ns
WI	$F(3,42) = 2.97$	*	.178	ns	ns	*	ns	ns	ns
H_SCALING		<i>P</i>	$\Omega^2$	1-2	1-3	1-4	2-3	2-4	3-4
SD	$F(3,43) = 2.37$	ns	.143	ns	ns	ns	ns	ns	ns
ZB_DF	$F(3,37) = 2.53$	ns	.169	ns	ns	ns	ns	ns	ns
ZB_IF	$F(3,31) = 5.27$	**	.339	ns	ns	ns	*	ns	**
RO	$F(3,30) = .78$	ns	.071	ns	ns	ns	ns	ns	ns
AM	$F(3,13) = 1.27$	ns	.223	ns	ns	ns	ns	ns	ns
GR	$F(3,54) = 2.59$	ns	.129	ns	ns	ns	ns	ns	ns
WI	$F(3,42) = 1.07$	ns	.069	ns	ns	ns	ns	ns	ns
L2_SCALING		<i>P</i>	$\Omega^2$	1-2	1-3	1-4	2-3	2-4	3-4
SD	$F(3,43) = 3.38$	*	.193	ns	ns	ns	ns	ns	ns
ZB_DF	$F(3,37) = 2.65$	ns	.177	ns	ns	ns	ns	ns	ns
ZB_IF	$F(3,31) = 6.00$	**	.372	ns	ns	**	ns	ns	ns
RO	$F(3,30) = 5.02$	**	.340	ns	**	*	ns	ns	ns
AM	$F(3,13) = 2.08$	ns	.324	ns	ns	ns	ns	ns	ns
GR	$F(3,54) = 1.33$	ns	.064	ns	ns	ns	ns	ns	ns
WI	$F(3,42) = 3.56$	*	.200	ns	ns	*	ns	ns	ns

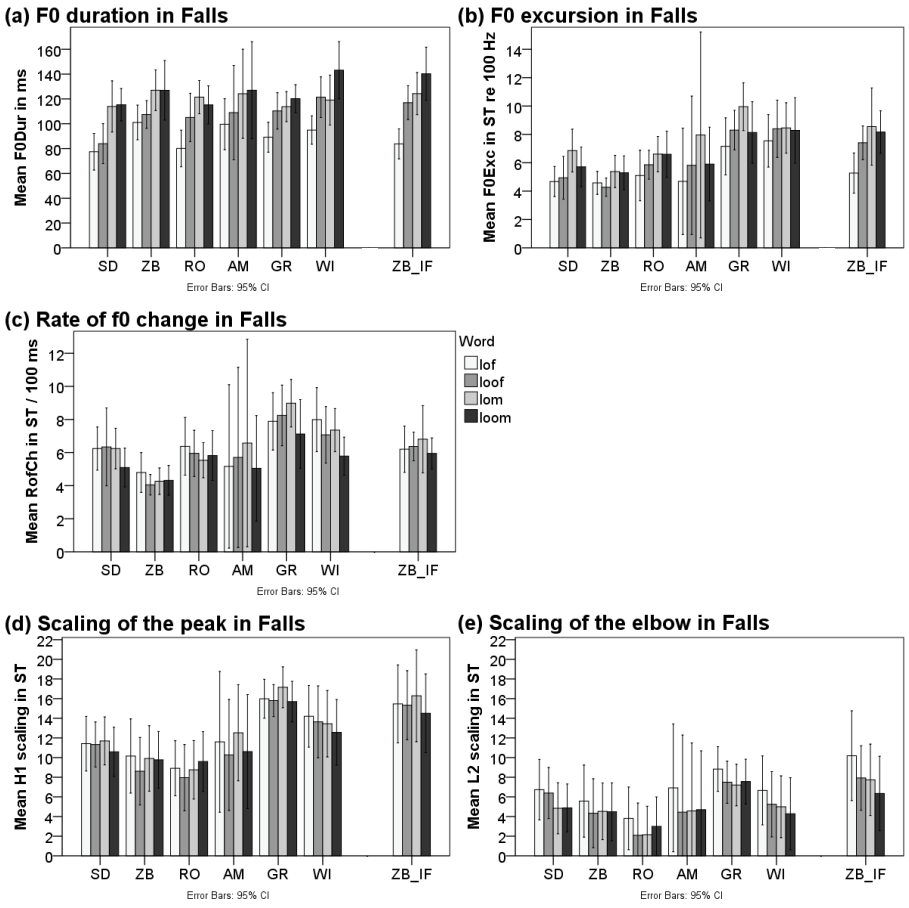


Figure 13. Mean f0 duration, f0 excursion, rate of f0 change, scaling of the nuclear peak (H\*) and scaling of the subsequent elbow (L2), broken down by word. Error bars represent 95% confidence interval.

### 3.3.5.2 Fall-rises

#### *Zuid-Beveland*

We report how time pressure affects the shape of fall-rises in Zuid-Beveland before presenting the results for the other varieties, because its realization is so specific for this variety that inclusion in a single analysis with data from the other varieties appeared irresponsible. A typical fall-rise (H\*L H%) falls from the nuclear peak before rising again at the end of the IP. However, a substantial proportion (37%) of the rhetorical questions in Zuid-Beveland was realized with a contour that can be described as *rising-rising*. The dip between the nuclear peak and the final high boundary tone is extremely small or absent. This is illustrated



in panels (a) and (b) of Figure 14, which compares Zuid-Beveland fall-rise and rise-rise contours.

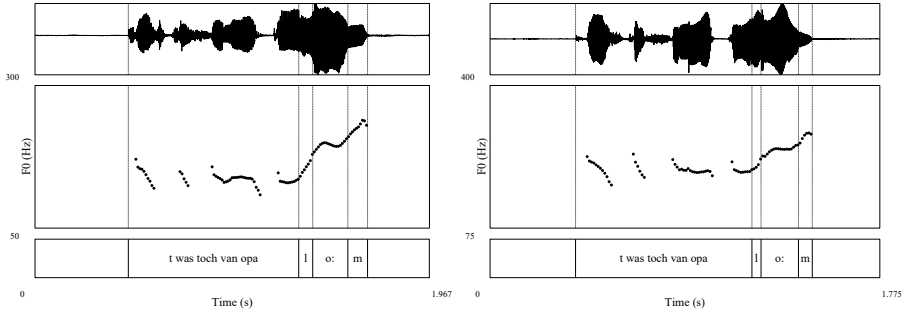


Figure 14. F0 tracks of a Zuid-Beveland fall-rise (panel a, male speaker) and rise-rise contours (panel b, female speaker) on the target word *loom* in ‘*t Was toch van opa Loom?*’ ‘But didn’t it belong to grandfather Loom?’.

The absence of a falling-rising movement in the rise-rises also means that it is not possible to determine the location of  $H^*$ . Consequently, the dependent variables used for the statistical analyses were calculated only on the basis of those Zuid-Beveland contours that contained a valley between two high targets. In this section we first present the results of the LME model for those data (based on data of 8 speakers). Next, we inspect averaged f0 contours of fall-rises (with and without all rise-rises) to describe the adjustments speakers make under time pressure.

Before carrying out analyses for the individual movements of the fall-rise, we checked whether f0 duration (from  $H^*$  to  $H^0$ ) was progressively shortened in shorter target words. We found an effect of WORD on total f0 duration, as table 16 illustrates. The table also shows a main effect of WORD on the f0 duration and the f0 excursion of the final rising part (part 2) as well as on the scaling of the valley. Pairwise comparisons show that the rising part in *lof* is significantly shorter and has a smaller excursion than *lom* and *loom*, and that the scaling of the valley is significantly higher in *lof* compared to *loom*. Generally, the bar charts in Figure 15 show that shorter words are realized with shorter falling f0 durations, smaller falling f0 excursions and slower falling rates of f0 change. These are all indications of undershoot of the low target, yet none of the dependent variables in the falling part of the contour (part 1) are affected by WORD. The strongest response to time pressure clearly takes place in the final movement.

Table 16. *Effects and pairwise comparisons for WORD on total f0 duration, f0 duration (F0DUR), f0 excursion (F0Exc), rate of f0 change (RofCh) of the falling (\_1) and rising part (\_2) separately, and on height of the peak (H\_Scaling), the subsequent valley (L2\_Scaling) and the final high boundary tone (H2\_Scaling).*

		<i>p</i>	$\Omega^2$	1-2	1-3	1-4	2-3	2-4	3-4
F0Dur	$F(3,19) = 10.06$	***	.634	ns	***	***	ns	ns	ns
F0Dur_1	$F(3,19) = 1.95$	ns	.252	ns	ns	ns	ns	ns	ns
F0Exc_1	$F(3,19) = 1.55$	ns	.209	ns	ns	ns	ns	ns	ns
RofCh_1	$F(3,19) = 1.19$	ns	.161	ns	ns	ns	ns	ns	ns
F0Dur_2	$F(3,18) = 6.12$	**	.524	ns	**	*	ns	ns	ns
F0Exc_2	$F(3,18) = 6.61$	**	.549	ns	**	**	ns	ns	ns
RofCh_2	$F(3,19) = 1.76$	ns	.237	ns	ns	ns	ns	ns	ns
H_Scaling	$F(3,18) = .83$	ns	.121	ns	ns	ns	ns	ns	ns
L2_Scaling	$F(3,18) = 4.31$	*	.419	ns	ns	*	ns	ns	ns
H2_Scaling	$F(3,18) = .83$	ns	.120	ns	ns	ns	ns	ns	ns

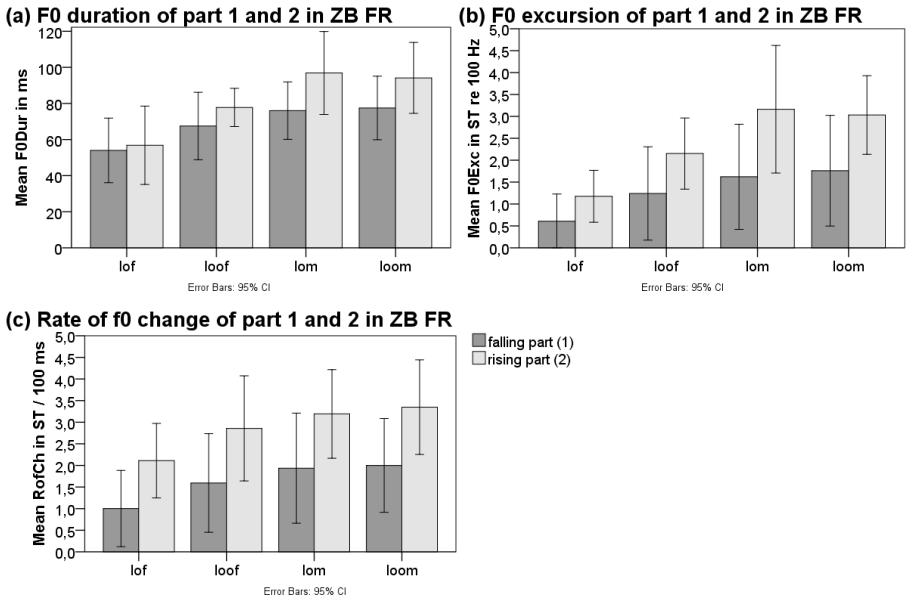


Figure 15. F0 duration in ms (panel a), f0 excursion in ST (panel b) and rate of f0 change in ST/100ms (panel c) of Part 1 and Part 2 of the fall-rise in Zuid-Beveland. Error bars represent the 95% confidence interval.

The averaged contours in panel (a) of Figure 16 (repeated from Figure 10) show that Zuid-Beveland H\*L H% contours are characterized by a drastic undershoot of the low target between the nuclear peak and the final boundary peak, with a tendency for more dramatic undershoot leading to rise-rises on words with shorter sonorant durations. The proportion of rise-rises within all fall-rises is 69% for *lof*, 35% for *loof*, 31% for *lom*, and 20% for *loom*. Besides undershoot of the low target, the contours in Figure 16 also suggest that speakers respond to time pressure by shortening (or truncating) both the falling and the final rising

movement, which they realize with a smaller and less steep  $f_0$  excursion on shorter words. If we add the *rise-rises* to the averaged contours as in panel (b), the same picture emerges, i.e. undershoot of the low target and truncation of the falling and particularly of the rising part of the contour.

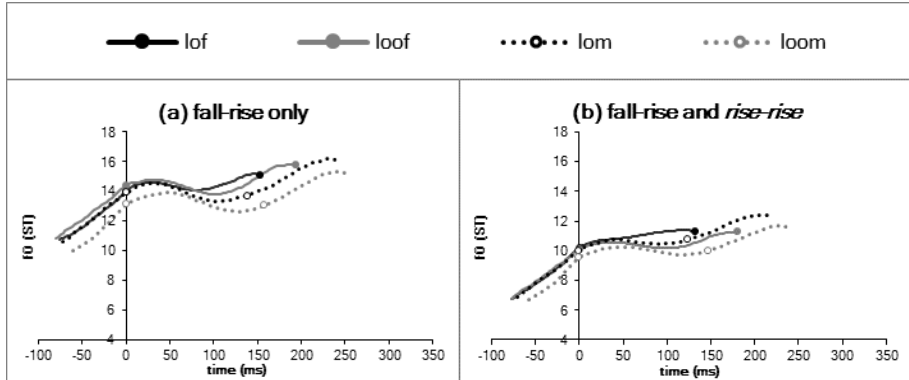


Figure 16. Averaged contours (in ST) of fall-rises *excluding* rise-rises (panel a) and fall-rises *including* rise-rises (panel b) on four target words in Zuid-Beveland. The beginning of the nuclear vowel is set to 0 ms. Circles represent the beginning and end of the nuclear vowel (/ɔ/ or /o:/).

### Other varieties

The total  $f_0$  duration of the falling-rising movement, ( $H^*$ -to- $H^0$ ) was subjected to statistical analysis. Table 17 shows a main effect of WORD in all varieties and almost all word pairs differed significantly in pairwise comparisons. Based on this result, we expect to find robust and regular adjustments of contours in response to time pressure in the realization of the fall-rise, either in the falling or in the rising part, if not in both.

First, we look at  $f_0$  duration of the falling and rising part of the  $H^*L H^0$  contour. In table 17, we see that WORD significantly affects both F0DUR\_1 and F0DUR\_2 in all varieties. Pairwise comparisons show that the shortest and longest word always differ significantly for both the falling and rising  $f_0$  duration, but that other adjacent word pairs differ considerably more often in the final rising part (L to  $H^0$ ). The exception to this general tendency is Winschoten, in which none of the adjacent word pairs (i.e., *lof-loof*, *loof-lom*, and *lom-loom*) differ significantly from each other in the final rising movement. The bar charts in Figure 17 confirm that the differences in  $f_0$  duration show a more regular pattern for L- $H^0$ , and that furthermore, the duration of the final rise decreases by a proportionally larger amount than the falling part, again with the exception of Winschoten.

Further down table 17 we can find the results for the effect of WORD on the shape of the falling part of  $H^*L H^0$ . We see that  $f_0$  excursion of the fall

(F0EXC\_1) is significantly affected in Standard Dutch, Amsterdam and Grou, and that we find a main effect of WORD on rate of change (ROCH\_1) in Standard Dutch, Amsterdam and Winschoten. The bar charts in Figure 18 show that, with the exception of *loof*, we observe the general pattern that f0 excursion of H\* to L is smaller and that the slope of these falls is fairly stable or less steep on shorter words. The pattern in Winschoten is strikingly the reverse, with shorter words having both larger excursions and steeper slopes, suggesting that these fall-rises are realized differently and that a different adjustment strategy is applied.

Moving on to the rising part of H\*L H%, table 17 shows that WORD significantly affects F0EXC\_2 in all varieties except Winschoten. More often than in the case of the fall, word pairs differ significantly from each other in pairwise comparisons. We only found a main effect of WORD on rate of change (RoCh\_2) in Standard Dutch, where pairwise comparisons show that the slope in *lof* is significantly less steep than in the other varieties. The bar charts in Figure 19 illustrate that speakers of Standard Dutch, Rotterdam, Amsterdam and Grou deal with the shape of the final rise in much the same way as they deal with the falling part. Shorter words have smaller f0 excursion and less steep slopes, which is indicative of truncation. Winschoten is different again, because both f0 excursion and slope of L to H% tend to remain stable across target words, showing that they do not compress the final rise as they do the falling movement.

Finally, table 17 gives the results for the scaling of tonal targets. The scaling of the nuclear peak was significantly affected by WORD in Standard Dutch, Rotterdam and Winschoten. Pairwise comparisons show that the effect was due to a high H\* on *lof* in Standard Dutch and Rotterdam, versus a low H\* in Winschoten. We also found an effect on the scaling of the final boundary tone in Standard Dutch, Rotterdam and Grou. As pairwise comparisons show, the effect was due to a high H% on *lom* in Grou and on *loom* in Standard Dutch. WORD had the strongest effect on scaling of the valley L, which was significant in all varieties except Winschoten. The valley was undershot and the falling movement truncated.

The mean f0 contours in Figure 10 above (middle panels) illustrate the effects of time pressure on the realization of fall-rises. We see, for example, undershoot of the valley in Standard Dutch, Rotterdam and Grou, and compression of the falling movement in Winschoten.

Table 17. *Effects and pairwise comparisons for WORD on total f0 duration, f0 duration (F0DUR), f0 excursion (F0EXC), rate of f0 change (ROFCH) of the falling (\_1) and rising part (\_2) separately, and on height of the peak (H1\_Scaling), the subsequent valley (L2\_Scaling) and the final high boundary tone (H2\_Scaling) in fall-rises.*

F0DUR		<i>p</i>	$\Omega^2$	1-2	1-3	1-4	2-3	2-4	3-4
SD	$F(3,57) = 59.24$	***	.787	***	***	***	ns	***	***
RO	$F(3,36) = 39.11$	***	.811	***	***	***	**	***	**
AM	$F(3,30) = 30.12$	***	.793	***	***	***	*	***	ns
GR	$F(3,59) = 61.80$	***	.808	***	***	***	***	***	***
WI	$F(3,41) = 19.72$	***	.602	**	***	***	ns	**	ns
F0DUR_1		<i>p</i>	$\Omega^2$	1-2	1-3	1-4	2-3	2-4	3-4
SD	$F(3,58) = 7.62$	***	.313	*	*	***	ns	ns	ns
RO	$F(3,36) = 7.91$	***	.407	*	ns	***	ns	ns	*
AM	$F(3,30) = 6.13$	**	.377	ns	**	**	ns	ns	ns
GR	$F(3,59) = 16.40$	***	.480	ns	**	***	ns	***	*
WI	$F(3,41) = 5.88$	**	.307	*	ns	**	ns	ns	ns
F0DUR_2			$\Omega^2$	1-2	1-3	1-4	2-3	2-4	3-4
SD	$F(3,57) = 55.38$	***	.791	***	***	***	ns	***	***
RO	$F(3,35) = 38.36$	***	.756	*	***	***	***	***	ns
AM	$F(3,41) = 22.37$	***	.621	*	***	***	*	***	ns
GR	$F(3,58) = 51.81$	***	.751	***	***	***	***	***	*
WI	$F(3,41) = 11.90$	***	.479	ns	***	***	ns	*	ns
F0EXC_1		<i>p</i>	$\Omega^2$	1-2	1-3	1-4	2-3	2-4	3-4
SD	$F(3,58) = 7.33$	***	.295	***	ns	***	ns	ns	ns
RO	$F(3,36) = 2.77$	ns	.192	ns	ns	ns	ns	ns	ns
AM	$F(3,30) = 3.48$	*	.264	*	ns	ns	ns	ns	ns
GR	$F(3,59) = 5.93$	***	.248	ns	ns	***	ns	ns	*
WI	$F(3,41) = 2.79$	ns	.166	ns	ns	ns	ns	*	ns
ROFCH_1		<i>p</i>	$\Omega^2$	1-2	1-3	1-4	2-3	2-4	3-4
SD	$F(3,58) = 4.03$	*	.179	*	ns	ns	ns	ns	ns
RO	$F(3,35) = 1.77$	ns	.131	ns	ns	ns	ns	ns	ns
AM	$F(3,31) = 2.97$	*	.235	ns	ns	ns	ns	ns	ns
GR	$F(3,59) = 2.59$	ns	.122	ns	ns	ns	ns	ns	ns
WI	$F(3,41) = 7.00$	***	.345	ns	ns	***	ns	*	*
F0EXC_2		<i>p</i>	$\Omega^2$	1-2	1-3	1-4	2-3	2-4	3-4
SD	$F(3,57) = 26.73$	***	.605	***	***	***	ns	**	**
RO	$F(3,35) = 16.90$	***	.674	**	***	***	ns	**	ns
AM	$F(3,30) = 6.99$	***	.445	**	*	***	ns	ns	ns
GR	$F(3,59) = 13.99$	***	.466	ns	***	***	*	**	ns
WI	$F(3,41) = 1.07$	ns	.069	ns	ns	ns	ns	ns	ns
ROFCH_2		<i>p</i>	$\Omega^2$	1-2	1-3	1-4	2-3	2-4	3-4
SD	$F(3,58) = 8.77$	***	.326	***	*	*	ns	ns	ns
RO	$F(3,36) = 2.21$	ns	.193	ns	ns	ns	ns	ns	ns
AM	$F(3,30) = 2.40$	ns	.197	ns	ns	ns	ns	ns	ns
GR	$F(3,60) = 1.65$	ns	.084	ns	ns	ns	ns	ns	ns
WI	$F(3,41) = 1.43$	ns	.091	ns	ns	ns	ns	ns	ns
H1_ST		<i>p</i>	$\Omega^2$	1-2	1-3	1-4	2-3	2-4	3-4
SD	$F(3,57) = 3.62$	*	.155	ns	*	ns	ns	ns	ns
RO	$F(3,35) = 3.60$	*	.235	ns	ns	*	ns	ns	ns
AM	$F(3,30) = .85$	ns	.079	ns	ns	ns	ns	ns	ns
GR	$F(3,59) = .76$	ns	.035	ns	ns	ns	ns	ns	ns
WI	$F(3,41) = 7.34$	***	.349	ns	ns	**	ns	**	**

L2_ST		<i>p</i>	$\Omega^2$	1-2	1-3	1-4	2-3	2-4	3-4
SD	$F(3,57) = 17.96$	***	.484	***	***	***	ns	ns	ns
RO	$F(3,35) = 15.14$	***	.566	***	**	***	ns	ns	*
AM	$F(3,30) = 3.86$	*	.283	*	ns	*	ns	ns	ns
GR	$F(3,59) = 16.63$	***	.462	***	ns	***	*	ns	***
WI	$F(3,41) = 1.58$	ns	.104	ns	ns	ns	ns	ns	ns
H2_ST		<i>p</i>	$\Omega^2$	1-2	1-3	1-4	2-3	2-4	3-4
SD	$F(3,57) = 4.09$	*	.181	ns	ns	**	ns	ns	ns
RO	$F(3,35) = 3.25$	*	.220	ns	ns	ns	ns	ns	ns
AM	$F(3,30) = 1.16$	ns	.101	ns	ns	ns	ns	ns	ns
GR	$F(3,59) = 8.04$	***	.301	ns	***	ns	***	ns	ns
WI	$F(3,41) = 2.28$	ns	.144	ns	ns	ns	ns	ns	ns

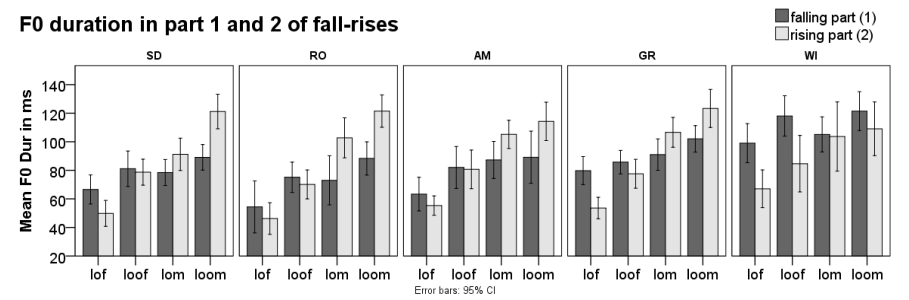


Figure 17. F0 duration from H\* to L (part 1) and L to H% (part 2) in ms, broken down by word and dialect. Error bars represent the 95% confidence interval.

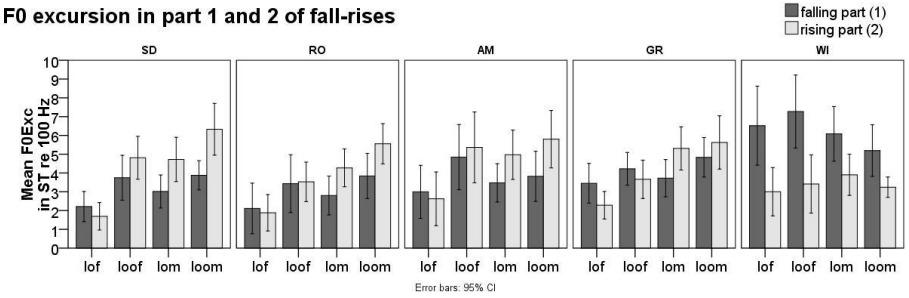


Figure 18. F0 excursion from H\* to L (part 1) and L to H% (part 2) in ST, broken down by word and dialect. Error bars represent the 95% confidence interval.

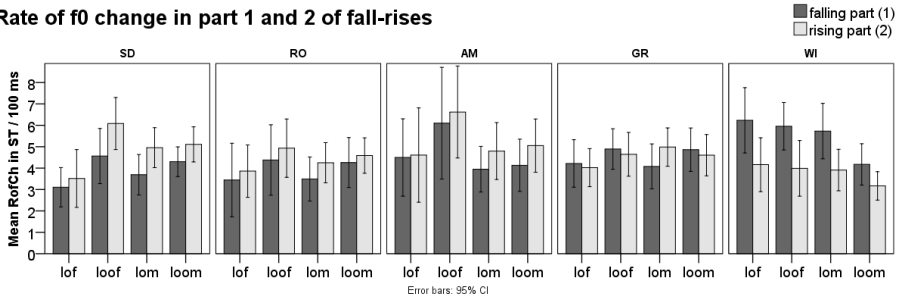
**Rate of f0 change in part 1 and 2 of fall-rises**

Figure 19. Rate of f0 change from H\* to L (part 1) and L to H% (part 2) in ST/100 ms, broken down by word and dialect. Error bars represent the 95% confidence interval.

### 3.3.5.3 Rises

The early start of the rise in *lof* and *lom* (see section 3.3.4.4) means that a relatively large proportion of the sonorant rime is available for the realization of the rise for these words. Accordingly, we do not expect to find systematic, step-wise variation between the four target words in terms of f0 duration, f0 excursion, rate of f0 change and scaling of the valley L and final H%.

For f0 duration, table 18 shows that WORD had a main effect in Standard Dutch, Rotterdam and Amsterdam. Pairwise comparisons for these varieties show that the rise from L to H tends to be longer on longer words, which is confirmed by the bar charts in Figure 20. Although still fairly small, differences between the words are somewhat larger than in the falls (3.3.5.1).

Next, table 18 shows that WORD does not affect F0EXC in any of the varieties except Standard Dutch, where f0 excursion in *lof* is significantly smaller than in *lom*. ROFCH is significantly affected by WORD in all varieties except Winschoten. The bar charts in Figure 20 suggest that the slope of the rise becomes steeper, the shorter the word. In other words, these speakers compress the rise, though not systematically from one word to the next.

Finally, we did not find an effect of WORD on the scaling of L and H except in Standard Dutch, where it is caused by a significantly higher H% in *lom* than in the other target words.

Thus, the main finding is that short-voweled rimes have earlier rise beginnings than long-voweled ones, but the shape of the rise doesn't vary greatly across words.

To conclude, despite the clear variation in timing of the rise between words with a short vowel and a long vowel, we found no such systematic variation in the shape of the rise.

Table 18. *Effects and pairwise comparisons for WORD on f0 duration (F0DUR), f0 excursion (F0EXC), rate of f0 change (ROFCH), L1\_Scaling and H1\_Scaling in rises.*

F0DUR		<i>p</i>	$\Omega^2$	1-2	1-3	1-4	2-3	2-4	3-4
SD	$F(3,24) = 11.51$	***	.607	ns	***	***	ns	*	ns
RO	$F(3,25) = 16.53$	***	.699	ns	***	***	*	**	ns
AM	$F(3,21) = 7.58$	***	.557	ns	**	*	*	ns	ns
GR	$F(3,13) = 1.03$	ns	.191	ns	ns	ns	ns	ns	ns
WI	$F(3,26) = 2.13$	ns	.118	ns	ns	ns	ns	ns	ns
F0EXC		<i>p</i>	$\Omega^2$	1-2	1-3	1-4	2-3	2-4	3-4
SD	$F(3,24) = 4.89$	**	.408	ns	**	ns	ns	ns	ns
RO	$F(3,25) = 2.63$	ns	.239	ns	ns	ns	ns	ns	ns
AM	$F(3,20) = 2.24$	ns	.257	ns	ns	ns	ns	ns	ns
GR	$F(3,13) = .24$	ns	.022	ns	ns	ns	ns	ns	ns
WI	$F(3,25) = 1.56$	ns	.155	ns	ns	ns	ns	ns	ns
ROFCH		<i>p</i>	$\Omega^2$	1-2	1-3	1-4	2-3	2-4	3-4
SD	$F(3,24) = 7.38$	***	.498	ns	ns	***	ns	**	*
RO	$F(3,25) = 9.50$	***	.537	ns	*	**	*	**	ns
AM	$F(3,20) = 9.14$	***	.585	ns	ns	***	ns	**	ns
GR	$F(3,13) = 5.04$	*	.538	*	ns	ns	ns	ns	ns
WI	$F(3,24) = .41$	ns	.983	ns	ns	ns	ns	ns	ns
L1_ST		<i>p</i>	$\Omega^2$	1-2	1-3	1-4	2-3	2-4	3-4
SD	$F(3,24) = 1.08$	ns	.112	ns	ns	ns	ns	ns	ns
RO	$F(3,24) = 1.47$	ns	.149	ns	ns	ns	ns	ns	ns
AM	$F(3,20) = 1.30$	ns	.164	ns	ns	ns	ns	ns	ns
GR	$F(3,13) = .86$	ns	.154	ns	ns	ns	ns	ns	ns
WI	$F(3,24) = .64$	ns	.076	ns	ns	ns	ns	ns	ns
H1_ST		<i>p</i>	$\Omega^2$	1-2	1-3	1-4	2-3	2-4	3-4
SD	$F(3,24) = 7.01$	***	.462	ns	***	ns	*	ns	*
RO	$F(3,24) = 1.54$	ns	.146	ns	ns	ns	ns	ns	ns
AM	$F(3,20) = 1.56$	ns	.187	ns	ns	ns	ns	ns	ns
GR	$F(3,14) = 1.99$	ns	.287	ns	ns	ns	ns	ns	ns
WI	$F(3,24) = 1.73$	ns	.183	ns	ns	ns	ns	ns	ns



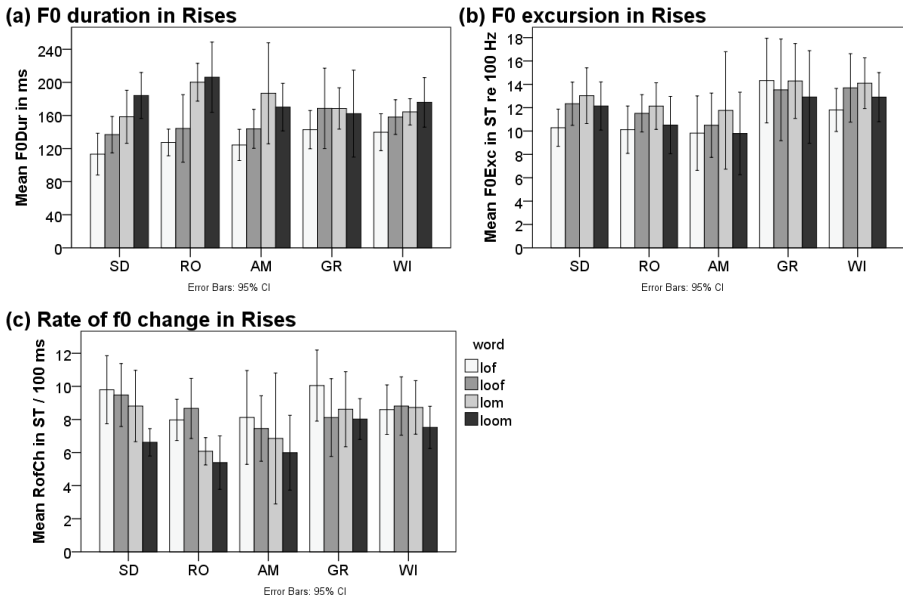


Figure 20. Averaged f0 duration, f0 excursion and rate of f0 change as measured from valley to peak, broken down by word.

### 3.4 Discussion

We have reported how speakers of six varieties of Dutch responded to time pressure caused by (1) complex vs simple contours on IP-final monosyllables bearing a nuclear pitch accent, and (2) a stepwise decrease in the duration of the sonorant rime of IP-final accented monosyllables. We looked at phonological responses (avoidance of complex contours), and at phonetic effects of time pressure on tonal timing, and on the shape of falls, fall-rises and rises.

#### 3.4.1 The realization of IP-final complex contours

In all varieties, speakers use the complex fall-rise on IP-final monosyllables in the rhetorical question condition in about 90% of the cases, while in the yes/no-questions, the fall-rise was used in about 25% of the data (though less so in Rotterdam and Zuid-Beveland). Whereas we cannot conclude that varieties of Dutch avoid complex contour IP-finally as in some varieties of German, we do see that its frequency of occurrence is affected by the availability of sonorant material.

We showed that in the central varieties (Standard Dutch, Rotterdam and Amsterdam) the proportion of complex contours is lower for shorter target words, both in the rhetorical questions and the yes/no-questions. Haan

(2002:109) reports that for Standard Dutch yes/no-questions with an *IP-medial* nuclear accent, the fall-rise (H\*L H%) is used in 72% of all read sentences. Seeing that the fall-rise is the preferred contour on non-final accented words in yes/no-questions, the low proportion of fall-rises on IP-final accented words in our data would appear to indicate that Standard Dutch speakers avoid the use of complex contours in this context.

We also provided evidence that speakers of all varieties facilitate the realization of the complex contour by increasing the duration of the sonorant material (compared to the simple falling contour) and by reducing the contour's pitch span, and that some varieties (Standard Dutch, Grou and Winschoten) place peaks earlier so as to leave a larger proportion of the sonorant rime available for the execution of the falling-rising melody. Zuid-Beveland was shown to be exceptional in its peak timing, with no difference in peak timing between fall-rises and falls at all. In this variety, the pitch span in the HLH contour is reduced so drastically that there is no need to take any other facilitating measures.

### 3.4.2 Peak / Valley retraction

#### Falls and fall-rises

For all varieties, we found a significant main effect of sonorant rime duration on the timing of nuclear peaks, measured from the beginning of the nuclear vowel, in H\*L L% and H\*L H%. Yet, closer inspection of the data suggests that, with the exception of Standard Dutch, the varieties do not retract the nuclear peak in a stepwise fashion for each shorter word, even though the peak in *lof* tends to occur earliest, and the peak in *loom* latest. The peak timing as a proportion of the sonorant rime shows that in falls the peak occurs about 20-30% into the sonorant rime, meaning that speakers leave roughly the same proportion of the sonorant rime available for the realization of the fall. In the case of the fall-rise, a weak effect is observed, indicating that speakers retract the peak in shorter rimes beyond what is predicted by a purely proportional peak timing.

#### Rises

The beginnings of rises varied significantly with target word. Yet, we only found evidence for Winschoten that the timing of the valley is actually affected by time pressure caused by decreasing sonorant rime duration. In this variety, the rise in the shortest word *lof* starts earlier than in the other target words. Interestingly, in all other varieties the rise started significantly earlier in words with a short vowel (*lof*, *lom*) than in words with a long vowel (*loof*, *loom*). This is illustrated by the averaged f0 contours of high and low rises in Figure 9 above (right-hand panel).

Since this effect did not follow from time pressure caused by decreasing sonorant rimes, we tested whether the target words' syllabic composition was responsible for the alignment differences. The results of the LME analysis in table 19, with

VOWEL (/ɔ/ vs. /o:/) and CODA (/f/ vs. /m/) as fixed factors and subject as random factor, showed that RELVALLEY significantly varies with VOWEL in all varieties except Winschoten. We did not find an effect of CODA, and we only found an interaction of VOWEL and CODA in Winschoten,  $F(1,26) = 5.59$ ,  $p < 0.05$ , which is due to the early valley alignment in *lof*, compared to the other target words.

Table 19. *Main effects of VOWEL and CODA on relative valley location (RELVALLEY) in rises.*

VOWEL		$p$	$\Omega^2$
SD	$F(1,24) = 11.32$	**	.449
RO	$F(1,25) = 10.18$	**	.301
AM	$F(1,22) = 22.17$	***	.489
GR	$F(1,13) = 36.82$	***	.762
WI	$F(1,27) = 3.11$	ns	.264
CODA		$p$	$\Omega^2$
SD	$F(1,24) = .25$	ns	.017
RO	$F(1,25) = .90$	ns	-.006
AM	$F(1,22) = .019$	ns	-.007
GR	$F(1,13) = 3.27$	ns	.231
WI	$F(1,27) = .46$	ns	.184

We will explore two possible explanations for the differences in relative valley alignment between short and long vowels in Standard Dutch, Rotterdam, Amsterdam and Grou. The first is that short vowels attract the use of the high rise over the low rise. The high rise has an early rise to the target of H\*, whereas the low rise starts later (Gussenhoven 2005). This behavior should be reflected in the proportion of high and low rises on each target word. Table 20 shows that except in Grou and Winschoten, *lof* and *lom* contain higher proportions of high rises than *loof* and *loom*.

Table 20. *Proportion of high rises in each target word.*

	<i>lof</i>	<i>loof</i>	<i>lom</i>	<i>loom</i>
SD	.35	.15	.36	.00
RO	.31	.08	.14	.00
AM	.40	.17	.43	.10
GR	.17	.63	.18	.20
WI	.00	.00	.00	.00

At the same time, the early rises in *lof* and *lom* could be due to a retraction of the valley in short vowels which is independent of the choice of low rise vs. high rise. We measured relative valley alignment of *low rises only* for the four target words in Standard Dutch, Rotterdam, Amsterdam, and Grou (i.e., the varieties that show the same behavior in Figure 8) to test this hypothesis. The results of a

mixed model analysis showed that VALLEYDELAY is significantly affected by VOWEL (/ɔ/ vs. /o:/;  $[F(1,93) = 55.76, p < .001, \Omega^2 = .537]$ ). RELVALLEY is significantly affected by VOWEL  $F(1,98) = 29.08, p < .001, \Omega^2 = .349$ ) and CODA ( $F(1,106) = 4.92, p < .05, \Omega^2 = .026$ ). Figure 21 illustrates the effect for VALLEYDELAY, and shows that rises in words with a short vowel are aligned earlier than in words with a long vowel, and also that rises in words with a sonorant coda are aligned earlier than in otherwise identical words with a fricative coda.

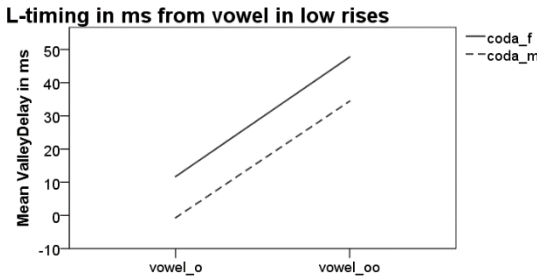


Figure 21. Valley alignment in low rises (L\*H H%), averaged across Standard Dutch, Rotterdam, Amsterdam and Grou (in ms from vowel), as a function of vowel length and coda type.

The results of the analysis for low rises shows that the early alignment in *lof* and *lom* cannot be explained by choice of contour alone. Even in the absence of high rises, rise beginnings in short vowels are aligned earlier. While this effect can readily be attributed to the difference in duration between long and short vowels<sup>10</sup>, the effect of the coda consonant is more mysterious. We currently have no explanation for this finding.

### 3.4.3 Truncation / Compression / Undershoot

#### Falls and rises

Our data show that in all varieties, the shape of the fall is not greatly affected by the duration of the sonorant portion. The duration of the steeply falling part of the contour (f0 duration of the fall) showed little variation, which is why differences in f0 excursion and rate of change of the contour are small between the target words. Those small differences were systematic enough to indicate that

<sup>10</sup> Ladd, Mennen, and Schepman (2000) and Ladd et al. (2009) showed that vowel class can have an effect on the alignment of pitch targets. They found a small effect on the alignment of the end of the prenuclear rise in Dutch, which was slightly earlier for a tense vowel than a lax vowel, despite their near-identical durations.

in response to time pressure, most varieties tended to truncate falls, but Winschoten tended to compress.

Adjustments in the shape of the rise in response to time pressure are small, but go in the direction of compression for all varieties, since the rate of change in shorter words is generally higher than in longer words.

### **Fall-rises**

Adjustments to time pressure can be most clearly observed in the complex contour, the fall-rise. We have shown that speakers make adjustments in both movements of the fall-rise. In Winschoten, however, speakers put more effort in making adjustments in the first part of the contour in order to accommodate the nuclear contour, than in the second part. The reverse is the case for the other varieties, which make adjustments in the second part more than the first. This is in agreement with the larger differences in  $f_0$  duration of the final rising movement, compared to the falling movement, that we saw between the target words of these varieties. In Winschoten, the shape of the falling movement is compressed in response to time pressure, whereas in the other varieties, we observed truncation of both the falling (= undershoot) and the rising movement of the fall-rise. The truncating response is stronger in some varieties than in others, and is stronger for the final rise than the falling part of the contour.

Undershoot of the valley is strongest in Standard Dutch and Zuid-Beveland. In Standard Dutch, we observed a reduced  $f_0$  excursion of the initial fall as well as a slower rate of change, corresponding to a raised valley. This could be seen in Figure 9, where the valleys of Grou, for instance are deeper. A quite different strategy is followed in Zuid-Beveland, where the severely undershot medial valleys are due to a manipulation of the shape of the fall towards a level pitch below the end pitch of the rising movement, in effect smothering out the falling part. The connection with time pressure is clear from the fact that this level rise or rise-rise contour is more frequently seen in the shorter words.

In terms of the categorization of responses, we find that speakers take shortcuts in both the falling and the rising part in Standard Dutch, Zuid-Beveland, Rotterdam, Amsterdam and Grou. Speakers of Winschoten increase their effort to realize the falling part, but not the rising part of the fall-rise.

### **3.5 Conclusion**

Our investigation of nuclear pitch contours in IP-final syllables in six non-tonal varieties of Dutch has yielded detailed phonetic data which allow a number of conclusions to be drawn.

First, all the theoretically possible ways of adjusting the realization of these pitch contours as set out in section 3.1 were attested in our data. The phonetic adjustments we found include (a) truncation, most clearly in the case of the nuclear fall and in the final rising movement of the nuclear fall-rise for Standard Dutch, Zuid-Beveland, Rotterdam, Amsterdam and Grou, and (b) compression in the case of the rise in all varieties, and in the nuclear fall and the falling part of the fall-rise in Winschoten. We have seen undershoot (c) of the low target in the fall-rise in Zuid-Beveland and Standard Dutch. Peak retraction (d) was attested in Standard Dutch, Rotterdam and Grou fall-rises. The comparison of complex and simple contours in section 3.3.3 illustrated (e) segmental lengthening in response to time pressure in all varieties. Finally, the data on the proportion of complex contours in yes/no-questions and rhetorical questions in section 3.3.2 provided evidence for (f) phonological adjustments in response to time pressure. Although complex contours are not categorically avoided in IP-final position, their frequency of occurrence is depressed in Standard Dutch, Rotterdam and Amsterdam. In short, these results show that both phonetic and phonological adjustments are variety-specific as well as contour-specific.

Clearly, it is not the case that a given dialect has one type of adjustment for a given contour. Frequently, there is a predominant adjustment (e.g., truncation of the rising part of the fall-rise in Standard Dutch) which is to some extent accompanied with another (e.g., peak retraction and undershoot for Standard Dutch fall-rises). Neither is it the case that we always see a stepwise increase in the use of these adjustments going from longer sonorant portions to shorter ones.

An unexpected result was the influence of vowel class on the alignment of rises. The rise is early in /lɔf/ and /lɔm/, and late in /lo:f/ and /lo:m/. That is, even though the sonorant rimes of /ɔm/ and /o:f/ have comparable durations, the beginning of the rise is determined by whether /ɔ/ or /o:/ occurs. Another unexpected result is the effect of the coda consonant on the alignment of the beginning of the rise, which is earlier in rimes containing /m/ than in those containing /f/. Future research should focus on these segmental effects.

Despite the dialect-specific, contour-specific, and context-specific nature of the responses to time pressure, the variation we found was sufficiently systematic that, to a large extent, these adjustments could in principle be incorporated into synthesis by rule programs for the generation of natural, dialect-specific speech.

Our findings represent a major advance over earlier investigations into responses to time pressure, which tended to interpret the observed responses in terms of only two mechanisms, truncation or compression. We have shown that it does not suffice to describe languages and contours as either truncating or compressing. Languages and varieties of the same language can resort to a variety of phonetic and phonological adjustment strategies, although some strategies are dominant. This conclusion is in partial agreement with that drawn by Ohl and

Pfützinger (2009), who say that “the choice of adjustment strategy is not a characteristic of an entire language system [...] or [...] variety [...]”. While there is variation in the choice of strategy-specific patterns and while choices are never absolute, we found clear dialect-specific preferences.

## References

- Bannert, R. and Bredvad, A. (1975). Temporal organisation of Swedish tonal accent: the effect of vowel duration. *Working papers 10, Phonetics Laboratory, Department of General Linguistics, Lund University, Sweden*, 1-36.
- Beckman, M.E. and Welby, P. (2006). *Elbow script*. Retrieved from <http://www.ling.ohio-state.edu/~welby/praat.html>.
- Boersma and Weenink (2008). *Praat: doing phonetics by computer* (Version 5.0.25) [computer program]. Retrieved May 31, 2008, from <http://www.praat.org/>.
- Caspers, J. and van Heuven, V.J. (1993). Effects of time pressure on the phonetic realization of the Dutch accent-lending pitch rise and fall. *Phonetica*, 50, 161-171.
- Del Giudice, A., Shosted, R., Davidson, K., Salihie, M., and Arvaniti, A. (2007). Comparing methods for locating pitch “elbows”. In: *Proceedings of the 16th international congress of phonetic sciences (ICPhS)*. Saarbrücken, Germany (pp. 1117-1120).
- Erikson, Y. and Alstermark, M. (1972). Fundamental frequency correlates of the grave word accent in Swedish: the effect of vowel duration. *STL, Quarterly Progress and Status Report*, 2-3, 53-60.
- Féry, C. (1993). *German intonational patterns*. Tübingen: Niemeyer.
- Gilles, P. (2005). *Regionale Prosodie im Deutschen: Variabilität in der Intonation von Abschluss und Weiterweisung*. Berlin: Walter de Gruyter.
- Grabe, E. (1998a). Comparative intonational phonology: English and German. *PhD thesis, MPI Series in Psycholinguistics*, 7. Nijmegen: Max Planck Institute for Psycholinguistics.
- Grabe, E. (1998b). Pitch accent realization in English and German. *Journal of Phonetics*, 26, 129-144.
- Grabe, E., Post, B., Nolan, F., and Farrar, K. (2000). Pitch accent realization in four varieties of British English. *Journal of Phonetics*, 28, 161-185.
- Gronnum (1989). Stress group patterns, sentence accents and sentence intonation in Southern Jutland (Sønderborg and Tønder) – with a view to German. *ARIPUC*, 23, 1-85.
- Gussenhoven, C. (2005). Transcription of Dutch intonation. In: Jun, S.-A. (Ed.), *Prosodic typology: The phonology of intonation and phrasing*. Oxford: Oxford University Press, 118-145.
- Gussenhoven, C. and van der Vliet, P. (1999). The phonology of tone and intonation in the Dutch dialect of Venlo. *Journal of Linguistics*, 35, 99-135.
- Haan, J. (2002). *Speaking of questions*. PhD thesis. Utrecht: LOT.
- Hanssen, J., Gussenhoven, C., and Peters, J. (submitted). Non-standard melodies and melody preferences in dialect-accented Dutch. *Submitted to Language & Speech*.
- Hanssen, J., Peters, J., and Gussenhoven, C. (2007). Phrase-final pitch accommodation effects in Dutch. In *Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS)*, Saarbrücken (pp. 1077-1080).
- Ladd, D.R. (2008). *Intonational phonology* (2nd ed.). Cambridge: Cambridge University Press.
- Ladd, D.R., Mennen, I., and Schepman, A. (2000). Phonological conditioning of peak alignment in rising pitch accents in Dutch. *Journal of the Acoustical Society of America*, 107, 2685-2696.
- Leben, W.R. (1976). The tones in English intonation. *Linguistic Analysis*, 2, 69-107.



- Lickley, R.J., Schepman, A., and Ladd, D.R. (2005). Alignment of "phrase accent" lows in Dutch falling rising questions: Theoretical and methodological implications. *Language and Speech*, 48: 157-183.
- Odé, C. (2005). Neutralization or truncation? The perception of two Russian pitch accents on utterance-final syllables. *Speech Communication*, 47, 71-79.
- Ohala, J.J. and Ewan, W.G. (1973). Speed of pitch change. *Journal of the Acoustical Society of America*, 53: 345(A).
- Ohl, C. and Pfitzinger, H. (2009). Compression and truncation revisited. In *INTERSPEECH* 2009, 2451-2454.
- Peters (1999). The timing of nuclear high accents in German dialects. In *Proceedings of the 14th International Congress of Phonetic Sciences (ICPhS)*, San Francisco, August 1999 (pp. 1877-1880).
- Peters, J. (2006). *Intonation deutscher Regionalsprachen*. Berlin: Walter de Gruyter.
- Petrone, C. and D'Imperio, M. (2009). Is tonal alignment interpretation independent of methodology? In: *INTERSPEECH* 2009, 2459-2462.
- Pierrehumbert, J.B. (1980). *The phonetics and phonology of English intonation*. PhD thesis, MIT.
- Pierrehumbert, J.B. (2000). Tonal elements and their alignment. In: Horne, M. (Ed.), *Prosody: Theory and Experiment. Studies Presented to Gosta Bruce*. Dordrecht: Kluwer, 11-26.
- Prieto, P., van Santen, J., and Hirschberg, J. (1995). Tonal alignment patterns in Spanish. *Journal of Phonetics*, 23, 429-451.
- Silverman, K. and Pierrehumbert, J.B. (1990). The timing of prenuclear high accents in English. In J. Kingston and M. Beckman (Eds.), *Papers in laboratory phonology I* (pp. 72-106). Cambridge: Cambridge University Press.
- Smakman, D. (2006). *Standard Dutch in the Netherlands. A sociolinguistic and phonetic description*. PhD thesis. Utrecht: LOT publications.
- Steele, S. (1986). Nuclear accent f0 peak location: Effect of rate, vowel, and number of syllables. *Journal of the Acoustical Society of America*, 80 (Suppl. 1), s51.
- Turk, A., Nakai, S., and Sugahara, M. (2006). Acoustic segment durations in prosodic research: a practical guide. In: Sudhoff, S., Lenertová, D., Meyer, R., Pappert, S., Augurzky, P., Mleinek, I., Richter, N. and Schliesser, J. (Eds), *Methods in empirical prosody research*. Berlin, New York: De Gruyter (= Language, Context, and Cognition, 3), 1-28.
- van den Berg, R., Gussenhoven, C., and Rietveld, T. (1992). Downstep in Dutch: Implications for a model. In G. Docherty and D.R. Ladd (Eds.), *Papers in Laboratory Phonology II: Gesture, Segment, Prosody* (pp. 335-359). Cambridge: Cambridge University Press.
- van de Ven, M. and Gussenhoven, C. (2011). On the timing of the final rise in Dutch falling-rising intonation contours. *Journal of Phonetics*, 39, 225-236.
- Waals, J. (1999). *An experimental view of the Dutch syllable*. PhD thesis, University of Utrecht. LOT-series 18. The Hague: Holland Academic Graphics.
- Xu, R. (2003). Measuring explained variation in linear mixed effects models. *Statistics in Medicine* 22, 3527-3541.
- Xu, Y. and Sun, X. (2002). Maximum speed of pitch change and how it may relate to speech. *Journal of the Acoustical Society of America*, 111, 1399-1413.
- Zhang, J. (2000). The phonetic basis for tonal melody mapping. In *Proceedings of the West Coast Conference on Formal Linguistics* 19. Somerville, MA: Cascadilla Press, 603-616.



### Abstract

This study examined the effects of focus on the realization of non-final nuclear falls in five varieties along the Dutch North-Sea coast. While phonetic effects surfaced more clearly in some varieties than others, we found no dialect-specific responses to the focus manipulation. In line with the findings for Standard Dutch reported in Hanssen, Peters, and Gussenhoven (2008), focus overall affected variables associated with the falling part of the nuclear contour. The results are interpreted in terms of hyper-articulation to express differences in communicative urgency. For sentences with higher degrees of urgency, speakers sought to maximize the pronunciation of the  $f_0$  fall inside the accented word, leading to shorter and steeper falls, which went down lower and sometimes started a little earlier. By lowering  $f_0$  in the postnuclear stretch even further, speakers added to the communicative effect of signaling greater urgency or importance in sentences with narrow or corrective focus, compared to broad focus.

### 4.1 Introduction

A focus constituent in West Germanic languages can be larger than the word carrying the nuclear pitch accent that signals the focus (e.g., Schmerling 1976, Gussenhoven 1983). While a distinction is traditionally made between broad focus (sentence-wide) and narrow focus (applying to constituents smaller than the sentence), to use the terms used in Ladd (1980), a focus constituent can have any size, including constituents smaller than the syllable when referred to metalinguistically (cf. van Heuven 1994). In (1a), the focus is ‘broad’, while in (1b) and (1c) the object NP is in focus. In addition to size, different ‘focus types’ have been distinguished. The focus meaning of (1a) and (1b) is ‘informational’ (Kiss 1998), while (1c) is ‘corrective’ (e.g. Gussenhoven 2007). This systematic focus ambiguity in *size* and *type* is illustrated in (1), where the focus constituent is indicated by square brackets.

---

<sup>1</sup> This chapter is a slightly revised version of: Hanssen, J., Peters, J., and Gussenhoven C. (2016). “Phonetic effects of focus in five varieties of Dutch”. In: *Proceedings of Speech Prosody 2016, Boston*, 736-740.

- |     |    |                           |                                                             |
|-----|----|---------------------------|-------------------------------------------------------------|
| (1) | a. | Broad<br>(informational)  | What's happening?<br>[They're drinking COFFEE].             |
|     | b. | Narrow<br>(informational) | What are they drinking?<br>They're drinking [COFFEE].       |
|     | c. | Narrow<br>(corrective)    | Are they drinking milk?<br>(No.) They're drinking [COFFEE]. |

While (1a,b,c) are generally analyzed as having the same phonological form, the difference in focus constituent size and focus type may have phonetic effects. Cross-linguistically, higher degrees of urgency or significance are associated with prominence-increasing properties, such as *higher* and *later or earlier peaks*; *larger*, *steeper* and *longer f<sub>0</sub> excursions*, and *longer segmental durations* (e.g., Eady and Cooper 1986, Xu 1999, Chen 2006, Smiljanić 2006, Chen and Gussenhoven 2008). Within West-Germanic, focus-related phonetic enhancement has been demonstrated for German (Féry and Kügler 2008, Baumann, Grice, and Steindamm 2006, Baumann et al. 2007, Peters 2002), Dutch (Hanssen, Peters, and Gussenhoven 2008, Peters, Hanssen, and Gussenhoven 2014) and to a lesser extent for English (Sityaev and House 2003).

There may be different motivations for speakers to enhance the perceivability of their speech. One is to promote that phonological contrasts are sufficiently distinctive, as happens in English when the laryngeal coda contrast is enhanced by vowel duration differences, as in the contrast between *seed* and *seat* or *strive* and *strife* (Stevens and Keyser 2010, Clements 2015). Another is immediately relevant to our topic, which is to promote the perception of meaning by pronouncing semantically significant morphemes more carefully (de Jong 1995, Ladd 2008). Since the phonological specification of intonational morphemes is notoriously localized, leaving sizeable stretches of speech tonally unspecified, the phonetic nature of enhancement may reveal where such specifications are located.

In a pilot experiment (Hanssen, Peters, and Gussenhoven 2008), we found that when a rising-falling nuclear pitch accent in Dutch is enhanced under intensification of the focus meaning, the enhancement is concentrated in the falling part. The low tone following the nuclear peak in narrow and corrective focus pitch accents was scaled lower and timed earlier, leading to a larger and steeper falling movement. The accent peak was timed slightly earlier but was otherwise unaffected. We found no effect of focus on the rise leading up to the nuclear peak. A small lengthening effect was found in the onset of the accented syllable. Effects were mainly between broad focus accents on the one hand and narrow and corrective focus accents on the other. Put differently, we found evidence for phonetic effects of focus *size*, but not focus *type*. The suggestion was therefore made that in Dutch the falling part of the nuclear accent is

communicatively more important, suggesting an off-ramp analysis of the accent (i.e., H\*L rather than L\*H).

The phonetic effects of focus found in other research can also be interpreted in terms of enhancement. However, most studies show enhancement in the nuclear peak and the part leading up to that peak, rather than the postnuclear stretch. In German, higher information values are associated with later and higher peaks, longer and larger rising pitch movements, and longer segmental durations (Baumann, Grice, and Steindamm 2006, Baumann et al. 2007). The same effects were demonstrated in Peters (2002) for narrow focus accents in non-final position, but in final position, narrow focus was mostly expressed through steeper falls after the nuclear peak. Since the nuclear accents in our pilot experiment were non-final, the effects on the falling part of the pitch accent are unexpected. They are also unexpected if we consider the results of Peters, Hanssen, and Gussenhoven (2014), which investigated the effect of differences in focus size smaller than the nuclear accented word in a number of varieties of Dutch and Low and High German. That investigation yielded no significant effect of focus size on the realization of the nuclear pitch accent in narrow corrective utterances. Compared to the baseline (wide information focus), however, corrective focus pitch accents were realized with increased segmental durations, higher and later peaks, lower preceding valleys, and larger rise and fall excursions. The specific enhancement strategies varied per dialect. Contrary to our pilot results, Peters, Hanssen, and Gussenhoven (2014) did not find peak retraction, nor lowering of postfocal  $f_0$ .

The purpose of this contribution is to see if our findings for Standard Dutch can be replicated in a number of dialects of Dutch. We test the hypothesis that the falling part of rising-falling nuclear pitch accents is hyperarticulated when it is communicatively more important. To this end, we designed a reading task with syntactically and lexically identical sentences that were phonologically ambiguous with respect to the size of the focus constituent (broad vs narrow) and focus meaning (i.e., informational vs. contrastive, more specifically corrective). The declarative sentences favored a rising-falling nuclear pitch accent on a non-final syllable.

## 4.2 Method

### 4.2.1 Varieties and subjects

We made recordings in five Dutch locations: Zeelandic in Zuid-Beveland (ZB), Hollandic in Rotterdam (RO) and Amsterdam (AM) (all Low Franconian), West Frisian in Grou (GR) and Low Saxon in Winschoten (WI). Data from 95 speakers were selected for analysis (17-23 speakers per variety, aged between 14-49, 40 male speakers). ZB, GR and WI speakers were bilingual with Dutch and their

local language. All regional speakers and their one or both parents were raised in the selected place and spoke the indigenous variety fluently. SD recordings were included if the geographical origin of the participants could not be determined by their accent.

#### 4.2.2 Materials and procedure

We used twelve declarative sentences of the type *We willen in Manderen blijven wonen* ('We want to stay in Manderen') as answers to a preceding question. Four questions elicited an answer with sentence-wide informational focus (henceforth broad focus, BF), four with narrow informational focus (NF) and four with narrow corrective focus (CF), assuming CF to have a higher 'information weight' than NF, and NF than BF. Example Q/A pairs are listed in (2) for each condition.

(2) **Broad focus**

- A. Wat is er met jullie? (*What's the matter?*)  
B. [We willen in Manderen blijven wonen.]

**Narrow focus**

- A. Waar willen jullie blijven wonen? (*Where do you want to stay?*)  
B. We willen in [Manderen] blijven wonen.

**Corrective focus**

- A. Willen jullie in Montfort blijven wonen? (*Do you want to stay in Montfort?*)  
B. Nee, we willen in [Manderen] blijven wonen.

A non-final falling nuclear pitch accent was expected to occur on the target word *Manderen*, a fictitious place name. Each of the target words had the metrical pattern sww (*Momberen*, *Memberen*, *Manderen* and *Munderen*) and was followed by two verbs with the pattern sw. The onset consonant was kept constant to be able to detect durational effects. We chose /m/ to limit interruptions or perturbations of the *f<sub>0</sub>* signal.

The Dutch set of test sentences was used for ZB, RO and AM. West Frisian (GR) and Low Saxon (WI) have their own standardized spelling systems and we therefore translated the Dutch materials into their local varieties, keeping the rhythmic, lexical, and segmental context as comparable as possible. Both language varieties reverse the order of the modal verb and the full verb, which means that the word following the test word is variable (e.g., *fytse* 'to cycle', *ite* 'to eat', *ride* 'to drive' for West Frisian).

Speakers were recorded in pairs and read each part of the mini-dialogue once. The recordings were made in a quiet room either in the homes of our speakers or in a public building, using a portable digital recorder (Zoom H4) with a 48kHz

sampling rate, 16 bit resolution and stereo format. The mini-dialogues were interspersed with 61 filler sentences (used for other experiments) and presented in a booklet in pseudo-randomized order, which was reversed for half of the subjects per variety.

To ensure that our subjects interpreted the information in the broad focus condition as all-new, these four sentences all appeared in the first block of approximately twenty sentences. NF sentences appeared in the second block and CF in the last block. In a control experiment with eight speakers of Standard Dutch, we statistically tested whether the order of presentation in block 1, 2 or 3 (ORDER) affected the phonetic realization of the nuclear pitch accent in terms of duration and  $f_0$ . We found that ORDER did not affect the timing and scaling of tonal targets, but shortened the segmental duration of words that were realized later in the reading task. In what follows, focus effects on segmental duration should therefore be interpreted with care if they are in the same direction as the effects of ORDER. In other words, if we find that CF or NF *shorten* segmental durations relative to NF or BF, respectively, this may be a consequence of ORDER instead of focus condition. Duration effects in the opposite direction (CF/NF showing *longer* durations compared to NF/BF) must be attributed to the factor FOCUS, although there is no way of establishing the exact degree to which differences are obscured by our confounding factor order of presentation.

#### 4.2.3 Data analysis

Using the speech processing software *Praat* (Boersma and Weenink 2008), we inserted the tonal and segmental labels listed in table 1, and stored their  $f_0$  value (f) and time (t) to compute the dependent variables in table 2. To neutralize gender differences in  $f_0$  excursion,  $f_0$  levels were converted to semitones re 100 Hz. Segmental labels were placed manually at segment boundaries on the basis of visual and auditory inspection of waveform and broadband spectrogram. Tonal labels were either low (L) or high (H). L1 and H were determined semi-automatically using a *Praat* function that traces the location of the highest or lowest  $f_0$  value in a selected interval. L2 was determined visually by selecting the location of the highest change in the speed of the  $f_0$  movement near the bottom line of the nuclear contour (cf. Hirst 2005). If two elbows were visible in the low-pitched section after the peak, we selected the first one.

Table 1. *Overview of tonal and segmental labels.*

<i>Tonal labels</i>	<i>description</i>
L1	elbow before nuclear peak
H	maximum $f_0$ of pitch accent
L2	elbow after nuclear peak
<i>Segmental labels</i>	<i>description</i>
O1	beginning of nuclear onset
V1	beginning of nuclear vowel
C1	beginning of nuclear coda
O2	beginning of onset first postnucl. unstressed syllable
O4	beginning of onset of first postnucl. stressed syllable
V4	beginning of vowel of first postnucl. stressed syllable

Table 2. *An overview of the dependent variables calculated from the tonal and segmental labels.*

Variables	Description	Formula
<i>Durational</i>		
ONSETDUR	duration of accented syllable onset (ms)	$t(V1) - t(O1)$
RIMEDUR	duration of accented syllable rime (ms)	$t(O2) - t(V1)$
SYLLDUR	duration of accented syllable (ms)	$t(O2) - t(O1)$
WORDDUR	duration of accented word (ms)	$t(O4) - t(O1)$
RISEDUR	duration of rise preceding nucl. peak (ms)	$t(H) - t(L1)$
RISEEXC	excursion of rise preceding nucl. peak (ST)	$f(H) - f(L1)$
RISESLOPE	rate of $f_0$ change of rise preceding nucl. peak (ST/s)	$RISEEXC/RISEDUR*1000$
<i>Shape</i>		
FALLDUR	duration of nuclear fall (ms)	$t(L2) - t(H)$
FALLXEC	excursion of nuclear fall (ST)	$f(L2) - f(H)$
FALLSLOPE	rate of $f_0$ change of nuclear fall (ST/s)	$FALLXEC/FALLDUR*1000$
POSTNEXC	excursion from nuclear peak to beginning of vowel of first postnuclear stressed syllable	$f(V4) - f(H)$
<i>Timing</i>		
L1_TIMING	distance of elbow prec. nucl. peak to beg. of onset (ms)	$t(L1) - t(O1)$
H_TIMING	distance of nucl.peak to beg. of nucl. vowel (ms)	$t(H) - t(V1)$
<i>Scaling</i>		
L1_SCALING	height of elbow preceding nucl.peak (ST re 100 Hz)	$f(L1)$
H_SCALING	height of nuclear peak (ST re 100 Hz)	$f(H)$
L2_SCALING	height of elbow following nucl. peak (ST re 100 Hz)	$f(L2)$
V4_SCALING	height at beginning of vowel of first postnucl. stressed syll. (ST re 100 Hz)	$f(V4)$

We analyzed the data using the Linear Mixed Effect Model procedure in SPSS, including SPEAKER and WORD as random factors, and FOCUS (BF/NF/CF) as fixed factor. Pairwise comparisons between the three levels of the fixed factor were carried out using the Bonferroni correction. To estimate the additional amount of variance explained by adding the fixed factor FOCUS to the model, as



opposed to a model that only includes the random factors, we used  $\Omega^2$ , following Xu (2003). The formula is

$$\Omega^2 = 1 - \frac{\text{variance residuals model random \& fixed}}{\text{variance residuals model random}}$$

To check the reliability of measurements, we compared the timing of the labels listed in table 1 for 72 sentences with a falling nuclear contour (12 items x 1 speaker x 6 varieties). Table 3 gives the mean difference between the timing of all segmental and tonal labels, as well as the inter-rater agreement, expressed as Cronbach's Alpha, which was at least 0.998 for all labels except L2<sup>2</sup> (Cronbach's Alpha 0.870).

Table 3. *Results of reliability test of the timing of 13 labels by 2 labelers.*

Label	Cronbach's Alpha	Mean absolute differences in ms
O1	0.999	4.26
N1	1.000	0.97
C1	1.000	1.49
O2	0.999	7.43
N2	1.000	-2.51
O4	0.999	3.12
N4	1.000	-1.54
O5	0.999	6.27
L1	0.998	-0.55
H1	0.999	1.24
L2	0.870	90.97

## 4.3 Results

### 4.3.1 Segmental duration

We observed a significant lengthening effect of FOCUS in Winschoten for ONSETDUR [ $F(2,159) = 3.70$ ,  $p < .05$ ,  $\Omega^2 = .0445$ ], with  $CF > BF$ ,  $p < .05$  in posthoc comparisons, and CODADUR [ $F(2,159) = 3.94$ ,  $p < .05$ ,  $\Omega^2 = .0468$ ], with  $NF > BF$ ,  $p < .05$ . FOCUS did not significantly affect segmental duration in any of the other varieties.

---

<sup>2</sup> The lower interrater agreement of L2 has two reasons. 1) the reliability test was carried out before a final adjustment of the labeling protocol for L2, described above; 2) in case of doubt, two L2 labels were placed (L2a and L2b), one of which was then selected by the first author. In the reliability test, label L2a is sometimes compared to L2b, which affects the outcome of that test.

### 4.3.2 Scaling of tonal targets

Table 4. *Estimated means of tonal scaling per variety.*

		BF	NF	CF		BF	NF	CF
L1	ZB	6.97	6.84	6.47	GR	9.54	9.64	9.73
H	ZB	12.26	11.89	11.77	GR	18.40	18.40	18.16
L2	ZB	6.68	6.49	6.24	GR	9.27	9.06	9.00
V4	ZB	6.55	6.12	5.98	GR	10.13	9.01	9.25
L1	RO	6.30	5.75	6.11	WI	8.77	8.52	7.81
H	RO	13.66	13.72	13.69	WI	15.86	16.33	16.02
L2	RO	5.64	5.37	5.32	WI	7.47	7.27	7.18
V4	RO	5.94	5.22	5.26	WI	7.31	6.86	6.71
L1	AM	5.03	5.01	4.70				
H	AM	12.99	12.84	11.95				
L2	AM	4.22	3.82	4.10				
V4	AM	7.39	6.54	5.86				

As table 4 shows, FOCUS generally had a lowering effect on low targets in all varieties and did not raise the high target of the nuclear peak, which was in fact lowered by FOCUS in ZB and AM.

The scaling of the elbow leading up to the nuclear peak (L1\_SCALING) was significantly affected in ZB, RO and WI, with BF higher than either CF or NF. (ZB [F(2,166) = 5.39,  $p < .01$   $\Omega^2 = .0609$ ] with BF > CF,  $p < .01$ ; RO [F(2,175) = 3.14,  $p < .05$   $\Omega^2 = .0329$ ] with BF > NF,  $p < .05$ ; WI [F(2,159) = 4.12],  $p < .05$   $\Omega^2 = 0.0525$  with BF > CF,  $p < .05$ .)

FOCUS had a significant lowering effect on peak height (H\_SCALING) in ZB [F(2,163) = 3.49,  $p < .05$   $\Omega^2 = .0438$ ], with BF > CF  $p < .05$  and in AM [F(2,180) = 9.00,  $p < .001$   $\Omega^2 = .0910$ ], with BF > CF,  $p < .001$  and NF > CF,  $p < .01$ .

We also found a lowering effect on the elbow after the peak (L2\_Scaling) in ZB [F(2,163) = 3.89,  $p < .05$   $\Omega^2 = .0486$ ], with BF higher than CF,  $p < .05$ .

Finally, the clearest effect of FOCUS on scaling was found when we looked at V4\_SCALING ( $f_0$  measured at the first postnuclear stressed vowel). FOCUS lowered postfocal material in all varieties, with BF higher than CF and/or NF. (ZB [F(2,164) = 3.26,  $p < .05$   $\Omega^2 = .0383$ ] with BF > CF,  $p < .05$ ; RO [F(2,178) = 5.97,  $p < .01$   $\Omega^2 = .0632$ ] with BF > NF,  $p < .01$  and BF > CF,  $p < .01$ ; AM [F(2,180) = 8.92,  $p < .001$   $\Omega^2 = .0902$ ] with BF > CF,  $p < .001$ ; GR [F(2,222) = 19.91,  $p < .001$   $\Omega^2 = .1517$ ] with BF > NF,  $p < .001$  and BF > CF,  $p < .001$ ; WI [F(2,159) = 5.18,  $p < .01$   $\Omega^2 = .0633$ ] with BF > CF,  $p < .01$ ).

### 4.3.3 Nuclear contour shape

This section looks at the shape (duration, excursion and slope) of the rise leading up to the nuclear peak, and the shape of the subsequent fall. We found no clear

pattern across varieties for the L1-to-H rise. The nuclear fall tended to be shorter in NF and CF compared to BF, with increasing excursions and steeper slopes. Shape differences between CF and NF were not always in the expected direction. We found significant effects of FOCUS on both rising and falling movements in RO, AM and WI. The effect is largest in AM, where FALLDUR is considerably shorter, while excursion is somewhat smaller, in corrective focus than in broad and narrow focus.

RISEDUR was affected in AM [ $F(2,184) = 4.37, p < .05, \Omega^2 = .0467$ ], with  $BF > CF, p < .05$ ; and in WI [ $F(2,159) = 6.78, p < .001, \Omega^2 = .0786$ ], with posthoc tests showing that  $BF < CF, p < .001$  and  $NF < CF, p < .05$ .

RISEEXC was affected in RO [ $F(2,175) = 3.22, p < .05, \Omega^2 = .0346$ ], with  $BF < NF, p < .05$ ; and in WI [ $F(2,161) = 3.40, p < .05, \Omega^2 = .0399$ ], with posthoc tests showing that  $BF < CF, p < .05$ .

FALLDUR was affected in RO [ $F(2,173) = 5.65, \Omega^2 = .0686$ ], with posthoc tests showing  $BF > NF, p < .05$  and  $BF > CF, p < .01$ ; and AM [ $F(2,183) = 6.284, p < .01, \Omega^2 = .0652$ ], with  $BF > CF, p < .01$  and  $NF > CF, p < .05$ .

FALLSLOPE was affected in AM [ $F(2,183) = 6.23, p < .01, \Omega^2 = .0644$ ], with  $BF > CF, p < .05$  and  $NF > CF, p < .01$ ; and WI [ $F(2,159) = 3.08, p < .05, \Omega^2 = .0406$ ], with no significant variation in posthoc tests.

FALLSLOPE was affected in RO [ $F(2,175) = 12.03, p < .001, \Omega^2 = .1255$ ], with  $BF < NF, p < .05$  and  $BF < CF, p < .001$ .

Furthermore, the excursion from H\* to the first postnuclear syllable (POSTNEXC) was significantly smaller in BF than in NF and/or CF in RO [ $F(2,176) = 4.17, p < .05, \Omega^2 = .0468$ ], with  $BF < NF, p < .05$  and  $BF < CF, p < .05$ ; GR [ $F(2,223) = 11.11, p < .001, \Omega^2 = .0922$ ], with  $BF < NF, p < .001$  and  $BF < CF, p < .05$ ; and WI [ $F(2,161) = 4.59, p < .05, \Omega^2 = .0540$ ], with  $BF < NF, p < .05$ .

The effect of focus on postnuclear excursion is illustrated in Figure 1.

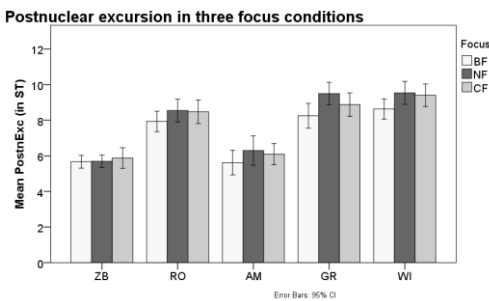


Figure 1. Postnuclear excursion (in ST) in broad focus, narrow focus and corrective focus. Error bars represent the 95% confidence interval.

4.3.4 Tonal timing

FOCUS had a significant effect on tonal timing in GR and WI, with timing of both L1 and H earlier in NF and/or CF, compared to broad focus sentences.

L1\_TIMING: GR  $[F(2,222) = 5.19, p < .01, \Omega^2 = .0431]$ , with  $BF > CF, p < .05$  and  $NF > CF, p < .05$ ; and WI  $[F(2,159) = 7.18, p < .001, \Omega^2 = .0872]$ , with  $BF > CF, p < .01$  and  $NF > CF, p < .01$ .

H\_TIMING: GR  $[F(2,222) = 8.71, p < .001, \Omega^2 = .0728]$ , with  $BF > NF, p < .01$  and  $BF > CF, p < .001$ ; and WI  $[F(2,158) = 8.31, p < .001, \Omega^2 = .0950]$ , with  $NF > CF, p < .001$ .

The effect of focus condition on peak timing in WI is different from the other varieties, because the peak is timed later, not earlier, in NF compared to BF. This finding is in line with the narrow focus results for H\_SCALING (section 4.3.3), although we have no explanation for it at present. This is also illustrated in Figure 2 below.

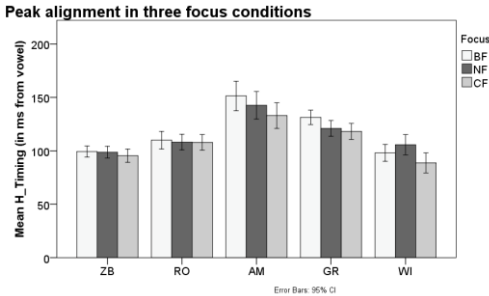


Figure 2. Peak timing (in ms from the nuclear vowel) in broad focus, narrow focus and corrective focus. Error bars represent the 95% confidence interval.

4.4 Discussion and conclusions

Our results show that in most of the varieties investigated, there are small differences in phonetic realization of nuclear falling contours as a function of focus condition. Segmental durations in WI were longer in the NF and CF condition than in the BF condition. Our ORDER confound may have obscured any other durational effects. Secondly, low targets were realized lower in all varieties in sentences with more intensified focus meanings. The lowering effect was most obvious in the  $f_0$  after the elbow, which means that speakers use the postnuclear stretch to express communicative differences. An additional lowering effect on the nuclear peak could be observed for ZB and AM. ZB was not otherwise affected by FOCUS. Timing effects were observed for WI and GR. Finally, the shape (duration, excursion and slope of  $f_0$  movements) was most

notably affected in RO, AM and WI. In AM, the effect of FOCUS on FALLDUR as well as RISEDUR reported in section 4.3.3 may incidentally have a phonological basis. The AM data includes both regular-peak falls and (late) half-completed falls. The latter are associated not only with later peaks, but also with shallower and longer falling movements. The longer rise and fall durations in BF sentences can be explained by a larger proportion of late-peak falls in this condition.

#### 4.4.1 Hyperarticulation

All effects of focus reported here are well-known from the literature summarized in section 4.1 and can be interpreted in terms of hyperarticulation (Lindblom 1990). Hyperarticulation can increase the prominence of (parts of) an utterance, with the purpose of increasing the distinctiveness of different levels of communicative urgency (Chen and Gussenhoven 2008). More prominence, for example in the form of larger pitch excursions, can be used to signal emphasis, enthusiasm or increased importance. Conversely, smaller pitch excursions are associated with less important information or a lack of interest.

Larger pitch excursions have been reported to go hand-in-hand with *higher* and *later* peaks. While this was confirmed in Peters, Hanssen, and Gussenhoven (2014), which is based on the same set of varieties and subjects as ours, our results show *earlier* peaks and *lowering* of  $f_0$  after the peak. Nevertheless, our results as well as those in Peters, Hanssen, and Gussenhoven (2014) can be interpreted as hyperarticulation. As described in Gussenhoven (2004), there are two ways in which a pitch peak can be enhanced. One is by raising it, a strategy which may evolve into peak delay as a substitute for raising, on the assumption that higher peaks are reached later. The other strategy is to hyperarticulate the pitch accent of which the peak is the realization. A more careful pronunciation of a falling pitch accent in an accented syllable may seek to maximize the pronunciation of the  $f_0$  fall inside the syllable rime, leading to a steeper fall that may begin earlier and reach lower (cf. Smiljanić 2006). The literature on West Germanic provides evidence for both these strategies. What unifies them is that they both serve to signal the communicative importance of the pitch accent's focus constituent. We have referred to this variation as communicative urgency, which has been manipulated by changing the focus meaning (focus type) and the size of the focus constituent, whereby smaller constituents are assumed to signal greater communicative urgency.

The results of our investigation tend to confirm the strategy used by speakers of Standard Dutch (Hanssen, Peters, and Gussenhoven 2008), whereby FOCUS overall affected variables associated with the falling part of the nuclear contour. Sentences with higher degrees of communicative urgency are expressed by steeper falls, which go down lower and may start a little earlier. The falls are also somewhat shorter as a result of the increased steepness that is sought by the

speaker. The results reported in Peters, Hanssen, and Gussenhoven (2014) point to the other strategy of later and higher peaks to maximize pitch excursions. We currently have no explanation for when which strategy is used. We note that the corrective focus test sentences in Peters, Hanssen, and Gussenhoven (2014) contained three levels of urgency (CF on the nuclear accented word, syllable or onset consonant), which may have triggered speakers to use increasingly higher peaks.

Just like Baumann et al. (2007), who found considerable speaker variation in the choice for particular strategies to mark focus structure, we also have not been able to identify variety-specific preferences for particular enhancement cues. Rather, the speaker's goal to hyperarticulate the fall can be attained by using a variety of strategies. We therefore support their suggestion that speakers can choose from different (phonological and phonetic) cues within a functional cluster (Local 2003) to mark focus structure.

#### **4.4.2 Focus size vs type**

Whereas the results for Standard Dutch reported in Hanssen, Peters, and Gussenhoven (2008) suggested that speakers enhance the pronunciation of the fall as a function of focus size rather than focus type, the current study of dialects fails to show that particular distinction. We found 30 significantly different BF-CF pairs (= variation in size and type), 13 significantly different BF-NF pairs (=size), and 12 significantly different NF-CF pairs (=type).

#### **4.4.3 Contextual clues**

It is likely that our speakers used, or even preferred, other cues besides differences in the realization of intonation structures to express or interpret focus structures. The literature reports, e.g., Downstep, deaccentuation, the use of a special pitch accent for focus, or the choice for and number of prenuclear accents. Other possibilities include visual cues (Swerts, Krahmer, and Avesani 2002), body language, eye contact, or the shared context between discourse partners.

One unintended contextual clue to information structure in our test materials was the presence of the focus marker 'no' at the start of the corrective focus sentence. This disambiguating morpheme may have had an effect on the necessity for speakers to express the focus structure phonetically. The fact that Baumann, Grice, and Steindamm (2006), whose test material also included the contrastive focus marker *nein* 'no', didn't find a durational difference between contrastive and non-contrastive focus, supports this possibility.

## References

- Baumann, S., Becker, J., Grice, M., and Mücke, D. (2007). Tonal and articulatory marking of focus in German. In: *Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS)*, Saarbrücken, 1029-1032
- Baumann, S., Grice, M., and Steindamm, S. (2006). Prosodic marking of focus domains – categorical or gradient? In: *Proceedings of Speech Prosody 2006, Dresden*, 301-304.
- Boersma, P. and Weenink, D. (2008). Praat: doing phonetics by computer (Version 5.0.25) [computer program]. Retrieved May 31, 2008, from <http://www.praat.org/>.
- Chen, Y. (2006). Durational adjustment under corrective focus in Standard Chinese. *Journal of Phonetics* 34, 176-201
- Chen, Y. and Gussenhoven, C. (2008). Emphasis and tonal implementation in Standard Chinese. *Journal of Phonetics* 36, 724-746.
- de Jong, K.J. (1995). The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *Journal of the Acoustical Society of America* 97, 491-504.
- Eady, S.J. and Cooper, W.E. (1986). Speech intonation and focus location in matched statements and questions. *Journal of the Acoustical Association of America* 80, 402-415.
- Féry, C. and Kügler, F. (2008). Pitch accent scaling on given, new and focused constituents in German. *Journal of Phonetics* 36, 680-703.
- Gussenhoven, C. (1983). Testing the reality of focus domains. *Language and Speech* 26, 61-80.
- Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge: Cambridge University Press.
- Gussenhoven, C. (2007). Types of focus in English. In: Lee, C., Gordon, M., Büring, D. (Eds.), *Topic and focus: Cross-linguistic perspectives on meaning and intonation*. Dordrecht: Springer, 83-100.
- Hanssen, J., Peters, J., and Gussenhoven, C. (2008). Prosodic effects of focus in Dutch declaratives. In: *Proceedings of Speech Prosody 2008, Campinas, Brazil*, 609-612.
- Hirst, D.J. (2005). Form and function in the representation of speech prosody. *Speech Communication* 46, 334-347.
- Kim, H. and Clements, G.N. (2015). The feature [tense]. In: Rialland, A., Ridouane, R., and van der Hulst, H. (Eds.), *Features in Phonology and Phonetics* (pp.159-178). Berlin/Boston: De Gruyter Mouton.
- Kiss, K.E. (1998). Identificational focus and information focus. *Language* 74, 245-273.
- Ladd, D.R. (1980). *The structure of intonational meaning. Evidence from English*. Bloomington: Indiana University Press.
- Ladd, D.R. (2008). *Intonational Phonology*. Cambridge: Cambridge University Press.
- Lindblom, B. (1990). Explaining phonetic variation: a sketch of the H & H theory. In: Hardcastle, W. and Marchal, A. (Eds.), *Speech production and speech modelling*. Dordrecht: Kluwer, 403-439.
- Local, J. (2003). Phonetics and talk-in-interaction. In: *Proceedings of the 15th International Conference of Phonetic Sciences (ICPhS)*, Barcelona, 115-118.
- Peters, J. (2002). Intonation und Fokus im Hamburgischen. *Linguistische Berichte* 189, 27-57.

- Peters, J., Hanssen, J., and Gussenhoven, C. (2014). The phonetic realization of focus in West Frisian, Low Saxon, High German, and three varieties of Dutch. In: *Journal of Phonetics* 46, 185-209.
- Schmerling, S. (1976). *Aspects of English sentence stress*. Austin: University of Texas Press.
- Sityaev, D. and House, J. (2003). Phonetic and phonological correlates of broad, narrow and contrastive focus in English. In: *Proceedings of the 15th International Conference of Phonetic Sciences (ICPhS)*, Barcelona, 1819-1822.
- Smiljanić, R. (2006). Early vs. late focus: Pitch-peak alignment in two dialects of Serbian and Croatian. In: Goldstein, L., Whalen, D.H., and Best, C.T. (Eds.), *Papers in Laboratory Phonology 8*. Berlin: Mouton de Gruyter, 495-518.
- Stevens, K.N. and Keyser, S.J. (2010). Quantal theory, enhancement and overlap. In: *Journal of Phonetics* 38, 10-19.
- Swerts, M., Krahmer, E., and Avesani, C. (2002). Prosodic marking of information status in Dutch and Italian: a comparative analysis. *Journal of Phonetics* 30, 629-654.
- van Heuven, V. (1994). What is the smallest prosodic domain? In: Keating, P. (Ed.), *Phonological structure and phonetic form. Papers in Laboratory Phonology III*. Cambridge: Cambridge University Press, 76-98.
- Xu, Y. (1999). Effects of tone and focus on the formation and alignment of  $f_0$  contours. *Journal of Phonetics* 27, 55-105.
- Xu, R. (2003). Measuring explained variation in linear mixed effects models. *Statistics in Medicine* 22, 3527-3541.



# FINAL AND NON-FINAL NUCLEAR CONTOURS ACROSS VARIETIES OF DUTCH, FRISIAN AND LOW SAXON<sup>1</sup>

---

## Chapter 5

### Abstract

This study describes the effect of gender, dialect and sentence-position on the realization of sentence-medial and sentence-final falling nuclear contours, and sentence final fall-rises. To this end, we examined production data from 119 speakers from six varieties of Dutch, Frisian and Low Saxon. As for gender, women have longer segmental durations than men, and produce a longer final rise in H\*L H% nuclear accents. The effect of sentence-position shows that compared to sentence-medial realizations, final nuclear contours are realized with longer segmental durations, earlier peaks, and shorter, smaller and steeper slopes. These effects can all be attributed to the lack of space to realize the contours at the end of an utterance, although not all dialects apply these adjustments in a uniform way. Regarding dialect effects, the most extreme realizational variation can be observed in the two geographically most extreme dialects. Southeastern (Zeelandic) falls are comparatively small and shallow, in sharp contrast with the large and steep Northeastern falls of Low Saxon. For most variables, the central varieties can generally be placed between the two extremes. However, they are different where peak timing is concerned. In non-final falls, speakers of these varieties often produce a late-peak fall which is particularly long and shallow. The main conclusion of the paper is that variation in the realization of nuclear contours in the Netherlands would seem to follow a geographical cline.

### 5.1 Introduction

The past fifteen years have shown an increased interest in the intonation of non-standard languages. In terms of Ladd's proposed taxonomy of cross-linguistic variation in intonation (Ladd 2008:116), differences can be *semantic*, *systemic*, *phonotactic* and *realizational*. A semantic difference could occur if a final rise is used only to signal non-finality in one dialect, but both non-finality and interrogativity in another. It could more generally be argued to be revealed by differences in the

---

<sup>1</sup> This chapter is a slightly revised version of: Hanssen, J., Gussenhoven C., and Peters, J. (forthcoming). "Final and non-final nuclear contours across varieties of Dutch Frisian and Low Saxon". To appear (2017) in: Zhang, H. (Ed.), *Prosodic Studies: Challenges and Prospects*.

frequencies with which the same phonological contours are used in different varieties, assuming that communicative needs are equal (Peters 2006). That is, dialects may show different preferences in the choice of nuclear contours in specific syntactic constructions or sentence types (Grabe and Post 2002, Grabe 2004, Gilles 2005, Dalton and Ní Chasaide 2007, Peters and Gussenhoven ms, Hanssen, Gussenhoven, and Peters ms). Systemic differences concern the tonal grammars (e.g., Gussenhoven and van der Vliet 1999, Kügler 2007, Prieto, D’Imperio, and Gili Fivela 2005, Peters 2006). Realizational differences concern the phonetic implementation of what are taken to be the same contours in different dialects, such as when nuclear contours at the end of the intonational phrase are truncated in one dialect but compressed in another (Grabe et al. 2000, Gilles 2005, Hanssen, Peters, and Gussenhoven ms). Or again, Dalton and Ní Chasaide (2005, 2007) and Kalaldehy, Dorn, and Ní Chasaide (2009) showed that some Irish dialects have fixed peak alignment, whereas others have variable peak alignment in contexts as a function of the number of unstressed syllables following the nuclear syllable. For varieties along the Dutch and German coast, Peters, Hanssen and Gussenhoven (2015) have documented that word boundary location of IP-medial words did not affect tonal timing. Other studies have investigated how focus position (final, non-final) or focus type (e.g., broad, narrow) affected the realization of intonation contours in non-standard varieties (e.g., O’Reilly, Dorn, and Ní Chasaide 2010, Avesani and Vayra 2003, Peters, Hanssen, and Gussenhoven 2014, and Hanssen, Peters, and Gussenhoven (2016).

The realization of intonation contours may also differ in more general ways instead of being governed by contextual factors such as segmental composition, upcoming word boundaries or focus condition. An example is peak timing in English, which is earlier than in Dutch and German (Atterer and Ladd 2004, Ladd et al. 2009). Attterer and Ladd (2004) and Mücke et al. (2009) reported that the peak in prenuclear and nuclear rising accents was timed later in Southern German than in Northern German. Kügler (2007) found that the peak in nuclear rising accents was timed earlier in the southern Swabian variety than in the eastern Upper Saxon variety of German. Dialectal variation in tonal timing has also been reported for varieties of Lowland Scots<sup>2</sup> (van Leyden 2004), German (Gilles 2005, Peters 1999), American English (Arvaniti and Garding 2007), Irish (Kalaldehy et al. 2009) and British English (Ladd et al. 2009).

Besides tonal timing, pitch excursion and overall pitch level can also show regional variation. Belgian women speak at a higher pitch than Dutch women

---

<sup>2</sup> The differences in tonal timing between the Lowland Scots varieties of Orkney and Shetland have an additional effect on *syllable duration*, which is longer for the Shetland variety. Van Leyden (2004:69) attributes the longer syllable duration to the fact that in Shetland, the entire rising movement is realized on the accented syllable, while in Orkney the peak of the rise is not realized until after the accented syllable.



(van Bezooijen 1993). Similarly, Gilles (2005:165) reported variation in  $f_0$  excursion of falling contours between speakers of eight varieties of German, with a maximum difference of more than 3 semitones (Duisburg vs. Dresden). Gilles suggested that these differences may be related to geography, with western varieties having larger pitch excursions than eastern varieties, which is confirmed by the finding of a 1.5 semitone difference between Saxon and Swabian German in Kügler (2007). Ulbrich (2004) reported differences in overall pitch range between speakers of two standard varieties of German (Swiss and Northern German)<sup>3</sup>. Finally, the dialects spoken on the Orkney and Shetland islands differ in overall pitch level, with intonation contours in the Orkney variety being realized at a higher pitch level (van Leyden 2004).

For Dutch, such dialectal characteristics had only been described informally, based on auditory impressions (van Es 1935, Daan 1938, Weijnen 1966). Recently, however, the results of two studies within the project *Intonation in Varieties of Dutch* have added support for the reality of a geographical cline of the sort suggested by Gilles (2005:165). Both studies compared a number of varieties spoken along the Dutch and German coast, covering Zeelandic and Hollandic Dutch, West Frisian, Dutch and German Low Saxon and Northern High German. First, Peters et al. (2014) investigated the effect of focus domain sizes smaller than the word on the realization of non-final falling nuclear contours. The second study (Peters, Hanssen, and Gussenhoven 2015) examined the effects of word boundary location on tonal timing of non-final nuclear falls. Besides reporting the results of the manipulations, the authors also described more general dialectal differences in contour realization. An inverted U-shaped pattern was observed for many phonetic variables. Overall, the ‘central’ varieties took more time to realize the  $f_0$  movements, which resulted in larger  $f_0$  excursions, higher peaks and later alignment of the pitch gesture with the segmental string. The accentual gestures of the peripheral varieties, on the other hand, were more compact, both in terms of duration and excursion. It can be concluded that the phonetic realization of nuclear falls in these varieties is determined more strongly by geographical proximity than by their linguistic grouping.

This paper looks at additional data from the project to see, first, whether we can replicate the finding of a geographical cline in the realization of non-final nuclear falling contours and, second, whether it is also found for IP-final nuclear contours (falls and fall-rises). Since the peripheral varieties of German Low Saxon and Northern High German are not included in our data set, we actually

---

<sup>3</sup> In addition, Ulbrich found differences in overall speech rate between the two varieties, which were caused by the number and duration of pauses within sentences.

expect to find only part of the inverted U-shape ‘’, roughly resembling a rising pattern ‘’. We focus on the effect of DIALECT on the phonetic realization of contours, as opposed to, for example, effects of time pressure, focus, or word boundary location. We will report dialectal differences in segmental duration as well as tonal timing, pitch excursion, pitch slope and overall pitch level. Finally, our data allow for a comparison of the regional differences in the realization of final and non-final nuclear contours. It is well-known that, compared to non-final falls, final falls may be realized with (1) longer segmental durations (final lengthening due to proximity of IP boundary, e.g., Wightman et al. 1992), (2) earlier nuclear peaks (e.g., Steele 1986, Prieto et al. 1995, Peters 1999), or (3) steeper or shorter falling excursions (e.g., Grabe 1998). It is less well-known, however, how dialects differ in their realization of final and non-final contours. Besides Standard Dutch, the comparison includes a Zeelandic dialect, two Hollandic varieties, West Frisian and Dutch Low Saxon. Our study excludes German Low Saxon and Northern High German, unlike Peters et al. (2014, 2015) reported above, which, in turn, did not look Dutch Low Saxon and Standard Dutch.

## 5.2 Procedure

### 5.2.1 Materials

We used three sets of sentences. The first set contained 4 declarative narrow focus carrier sentences with a non-final falling pitch accent (nf-FALL); the second set contained 4 declarative narrow focus carrier sentences with an IP-final falling pitch accent (f-FALL); and the last set contained 4 rhetorical questions with an IP-final falling-rising pitch accent (f-FR). All 12 carrier sentences (labeled ‘B’) were preceded by a context sentence (‘A’) with which they formed a mini-dialogue, as illustrated in table 1. In the non-final declaratives, the target words consisted of fictitious place names *Momberen*, *Memberen*, *Manderen* and *Munderen*<sup>4</sup>, which had the metrical pattern *swm*, in which the segmental structure of the accentable first syllable was Nasal-V-Nasal, followed by a voiced plosive onset consonant. They were followed by two verbs with the pattern *sw*. In the carrier sentences for the accentable IP-final position, four fictitious monosyllabic proper names, *Lof*, *Loof*, *Lom*, *Loom*, were used as target words in each pragmatic condition. These varied in the rime only, where short [ɔ] and long [o:] combined with voiceless [f] and sonorant [m].

---

<sup>4</sup> Speakers of Standard Dutch produced three sentences each, with the target words *Manderen*, *Bunderen* and *Lunteren*.

Table 1. *Dutch context sentences and experimental sentences used to elicit non-final falls, final falls and final fall-rises, with English translations. The target sentences are printed in bold; the word carrying the nuclear pitch accent is capitalized.*

	Context sentence	Carrier sentence
nf-FALL	Waar zouden je oom en tante willen wonen? <i>Where would your uncle and aunt want to live?</i>	<b>Ze zouden bij MANDEREN willen wonen.</b> <i>They'd like to live near Manderen.</i>
f-FALL	Met wie gaat je baas morgen trouwen? <i>Who will your boss marry tomorrow?</i>	<b>Hij trouwt met mevrouw de LOOM.</b> <i>He'll marry Mrs. De Loom.</i>
f-FR	Dit antieke horloge is nog van opa Thijssen geweest. <i>This antique wristwatch used to belong to grandfather Thijssen.</i>	<b>Het was toch van opa LOOM?</b> <i>But didn't it belong to grandfather Loom?</i>

We collected the Standard Dutch data in a pilot experiment. A slightly modified version of the pilot sentences formed the Dutch set of experimental sentences, which was used for Zuid-Beveland, Rotterdam and Amsterdam. Speakers from Zuid-Beveland translated the Dutch sentences to their variety as they went along. We translated the sentences into the local language for speakers from West Frisian and Low Saxon, who have their own standardized spelling system. For all varieties, the rhythmic, lexical, and segmental context was kept comparable to the Standard Dutch materials as much as possible. An overview of the sentences in all language versions is given in the Appendix.

### 5.2.2 Varieties and subjects

Recordings were made in five locations along the Dutch coast, covering four dialect groups (Figure 1). Zeelandic Dutch in Zuid-Beveland (ZB), Southern Hollandic in Rotterdam (RO), and Northern Hollandic in Amsterdam (AM) all belong to the Low Franconian language family. We also recorded a West Frisian in Grou (GR) and a Low Saxon variety in Winschoten (WI). The Standard Dutch (SD) speakers were recorded in Nijmegen. Historically, Standard Dutch has close relations to western varieties such as Rotterdam and Amsterdam (cf. Smakman 2006 and references therein).



Figure 1. Recording locations in the Netherlands.

We recorded 119 speakers (between 18 and 23 speakers for each variety), 49 of whom were male. They were aged between 14 and 49. Participants were university students (SD), secondary school students (ZB), members of a Scouting club (RO, AM) or members of the local community (GR, WI). The speakers from Zuid-Beveland, Grou, and Winschoten were bilingual with Standard Dutch and their local language. All regional speakers and at least one of their parents were raised in the selected place and spoke the indigenous variety fluently. For Standard Dutch, the procedure was different, as the area where this variety is spoken is less determined by geographical boundaries. Speakers could participate if they reported to speak Standard Dutch. Besides self-reporting, two Dutch phoneticians independently judged each recording. Recordings were included if the judges agreed that the geographical and linguistic origin of the participants could not be determined by their accent. Except for the speakers of West Frisian and Standard Dutch, our speakers were less familiar with their local language as a written language, which may have had a negative influence on the fluency of the speech in the reading task of some speakers.

Participants' recordings were excluded if they were (highly) disfluent or appeared to the experimenter not to speak naturally, if the speakers afterwards reported that they were dyslectic or had hearing problems, or if the speakers turned out not to satisfy the requirements with respect to their linguistic and/or geographical background. All participants were naive as to the purpose of the task and were paid for their participation.

### 5.2.3 Recording procedure and data selection

To avoid listing effects, the 12 mini-dialogues were interspersed with 61 filler sentences (used for other experiments) and presented in a booklet, in randomized order which was reversed for half of the subjects per variety. Speakers were recorded in pairs to reduce any effects of the experimenter's presence and the nature of the task on their dialect level. One speaker read the context sentence and the other the carrier sentence. The participants switched roles at the end of the task after they had repeated any mispronounced sentences.

The Standard Dutch recordings were made in a professional studio at Radboud University Nijmegen; recordings of the local varieties were made in a quiet room either in the homes of our speakers or in a public building. We used a portable digital recorder (Tascam HD P2 for Standard Dutch and Zoom H4 for all other varieties) with a 48 kHz sampling rate, 16 bit resolution and stereo format. The participants wore head-mounted Shure WH30XLR or Sennheiser MKE 2 wired condenser microphones.

All recorded target sentences were converted to monaural files and stored on computer disk as separate wav files with a sampling rate of 48 kHz and 16 bit resolution. Utterances were excluded from further analysis if they showed deviant pitch patterns due to accent position or choice of nuclear pitch contour. More specifically, for the declarative condition, we only included utterances that were realized with a nuclear falling contour (H\*L L%), and only utterances with a fall-rise (H\*L H%) were selected for the rhetorical questions. Zuid-Beveland speakers often realized the fall-rise as a 'rise-rise', that is, a sequence of rising movements without a low turning point between the two peaks<sup>5</sup>. Therefore, we only included the eight ZB participants whose data could be labeled as H\*L H%, i.e. as fall-rises with a low turning point. A final remark with respect to data selection is that speakers of Winschoten pronounced the trisyllabic target word in non-final falls 'Manderen' as disyllabic [mandə:n] in over 70% of the cases, whereas in other varieties it was realized with three syllables, [mandərə]. We nevertheless included Winschoten in our analyses, and will interpret the results in this context.

The total number of speakers whose data was used for analysis is given in table 2, broken down by variety, sentence condition and gender.

---

<sup>5</sup> The shape of the Amsterdam falls, and the shape of the Zuid-Beveland 'rise-rise' are discussed in Hanssen, Gussenhoven, and Peters (ms), along with their phonological interpretations.

Table 2. *Number of speakers used in the analyses, broken down by variety, sentence condition and gender.*

<i>nf-</i> <i>FALL</i>				<i>f-</i> <i>FALL</i>			<i>f</i> -FR		
	<i>F</i>	<i>M</i>	<i>total</i>	<i>F</i>	<i>M</i>	<i>total</i>	<i>F</i>	<i>M</i>	<i>total</i>
SD	13	8	21	9	8	17	13	9	22
ZB	7	10	17	7	8	15	6	2	8
RO	7	12	19	3	10	13	7	8	15
AM	7	11	18	4	2	6	6	6	12
GR	20	3	23	18	3	21	20	2	22
WI	13	4	17	12	4	16	13	2	15
<i>total</i>	67	48	115	53	35	88	65	29	94

### 5.2.4 Variables and analysis

Acoustic and auditory analysis of the data was done with the help of the speech processing software package *Praat* (Boersma and Weenink 2008). We inserted the labels listed in table 3 and stored their time (t) and  $f_0$  value (f), which was converted from Hz to semitones (ST re 100 Hz).

Segmental labels were all placed manually at segment boundaries. The boundaries were determined according to general practice, on the basis of visual inspection of waveform and broadband spectrogram, aided by auditory information (Turk, Nakai, and Sugahara 2006). We placed all labels at negative-to-positive zero-crossings.

Tonal labels were either low (L) or high (H). All high tones (H and H2), plus all low tones in fall-rises, were determined semi-automatically using a Praat function that traces the location of the highest or lowest  $f_0$  value in a selected interval. Determining the location of the elbow after the nuclear peak (L) in final and non-final falls was less straightforward, especially in those cases where contours displayed a gradual change in slope (cf. Del Giudice et al. 2007; Petrone and D’Imperio 2009). To increase interrater agreement, we therefore determined L in falls visually by selecting the location of the highest change in the speed of the  $f_0$  movement near the bottom line of the nuclear contour<sup>6</sup>. If two elbows were visible in the low-pitched section after the peak, we selected the first one. Each label was checked and corrected for tracking errors due to pitch perturbations.

Using the labels in table 3, we then computed the dependent variables listed in table 4.

---

<sup>6</sup> This point roughly corresponds to the location in the  $f_0$  curve where the first derivative (or slope) is zero. As such, our method is the manual version of the MOMEL algorithm used as input for the INTSINT transcription system for the representation of intonation (cf. Hirst 2005).



Table 3. *Overview of acoustic measurement labels.*

<i>Pitch targets</i>		<i>nf-</i>	<i>f-</i>	<i>f-</i>
		<i>FALL</i>	<i>FALL</i>	<i>FR</i>
H	maximum f0 of nuclear pitch accent (nuclear peak)	✓	✓	✓
L	elbow after nuclear peak	✓	✓	✓
H2	Maximum f0 of final boundary tone			✓
<i>Segmental boundaries</i>				
O1	beginning of onset of nuclear syllable	✓	✓	✓
N1	beginning of rime of nuclear syllable	✓	✓	✓
C1	beginning of coda of nuclear syllable	✓	✓	✓
O2	end of rime of nuclear syllable	✓	✓	✓

Table 4. *Acoustic variables used in the comparison of non-final final nuclear contours in five varieties. A tick in the row below the three contour names indicates this variable was measured for that contour.*

<i>Durational variables</i>		<i>formula</i>	<i>nf-</i>	<i>f-</i>	<i>f-</i>
			<i>FALL</i>	<i>FALL</i>	<i>FR</i>
SONRIME	the duration of the sonorant rime of the nuclear syllable in ms	$t(O2) - t(N1)$	✓	✓	✓
<i>Timing variables</i>					
RELPEAK	the timing of H as a proportion of the sonorant rime duration in %	$\frac{(t(H) - t(N1))}{(t(O2) - t(N1))} * 100$	✓	✓	✓
<i>Scaling variables</i>					
H-SCALING	the height of the nuclear peak in ST re 100 Hz	$f(H)$	✓	✓	✓
L-SCALING	the height of the elbow following the nuclear peak in ST re 100 Hz	$f(L)$	✓	✓	✓
H2-SCALING	the height of the final boundary tone in fall-rises in ST re 100 Hz	$f(H2)$			✓
<i>Contour shape variables</i>					
FALLDURATION	the duration of the fall following the nuclear peak in ms	$t(L) - t(H)$	✓	✓	✓
FALLEXCURSION	the excursion of the fall following the nuclear peak in ST	$f(L) - f(H)$	✓	✓	✓
FALLSLOPE	the rate of change of the fall following the nuclear peak in ST/s	$\frac{FALLEXCURSION}{FALLDURATION} * 1000$	✓	✓	✓
RISEDURATION	the duration of the final rise in fall-rises in ms	$t(H2) - t(L)$			✓
RISEEXCURSION	the excursion of the final rise in fall-rises in ST	$f(H2) - f(L)$			✓

RISESLOPE	the rate of change of final rise in fall-rises in ST/s	RISEEXCURSION/ RISEDURATION *1000	✓
RATIOFRDUR	relation between duration of falling and rising part of fall-rise	FALLDURATION / RISEDURATION	✓
RATIOFREXC	relation between excursion of falling and rising part of fall-rise	FALLEXCURSION / RISEEXCURSION	✓
RATIOFRSLOPE	relation between slope of falling and rising part of fall-rise	FALLSLOPE / RISESLOPE	✓

Unless otherwise stated, we analyzed the data using the Linear Mixed Effects Model in SPSS, including SPEAKER and SENTENCE as random factors, and DIALECT (SD, ZB, RO, AM, GR, WI) and GENDER as fixed factors. SENTENCE\_CONDITION (nf-FALL, f-FALL, f-FR) was included as a fixed factor in the model for those dependent variables that were measured for all contours. Pairwise comparisons between the levels of the fixed factor were carried out using the Bonferroni correction.

Since female speakers on average speak at a higher pitch level than male speakers (225 Hz vs. 125 Hz), we measured f0 in semitones. This will to a large extent normalize gender variation where excursion sizes are concerned, but will not normalize differences in scaling of individual pitch targets (such as the scaling of the nuclear peak). The effects of DIALECT on tonal scaling (H-SCALING, L-SCALING, H2-SCALING) will be reported only for female speakers, since the number of male speakers is comparatively small for some contour conditions in some varieties (see table 2).

5.3 Results

5.3.1 Sonorant rime duration

The bar chart<sup>7</sup> in Figure 2, which gives sonorant rime durations by contour type for varieties separately, allows us to make two observations. First, sonorant rime duration increases from nf-FALLS, to f-FALLS and f-FR. This pattern holds across all dialects. Second, rime durations tend to gradually increase from the southwest (ZB) to the northeast (WI).

<sup>7</sup> The bars representing SD are separated from the other bars throughout the paper to underline the fact that, unlike the local varieties, SD speakers do not form a geographically coherent group.

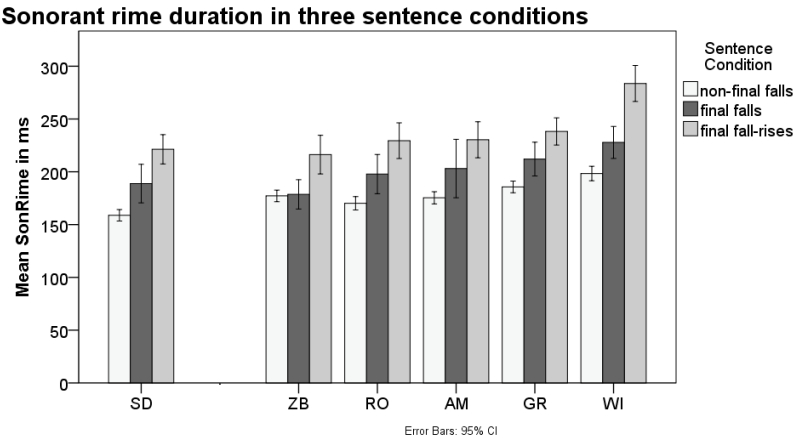


Figure 2. Mean sonorant rime duration in non-final falls, final falls and final fall-rises for each variety. Error bars represent the 95% confidence interval.

We found main effects of DIALECT, GENDER and SENTENCE\_CONDITION on the duration of the sonorant rime, and interactions between DIALECT x SENTENCE\_CONDITION, and GENDER x SENTENCE\_CONDITION.

Table 5. *Effect of DIALECT, GENDER and SENTENCE\_CONDITION on SONRIME.*

<i>Dialect</i>	$F(5,111) = 3.83$	$p < .01$
<i>Gender</i>	$F(1,110) = 22.16$	$p < .001$
<i>Sentence_condition</i>	$F(2,10) = 62.86$	$p < .001$
<i>Dialect x Sentence_condition</i>	$F(10,957) = 3.48$	$p < .001$
<i>Gender x Sentence_condition</i>	$F(2,964) = 6.96$	$p < .001$

Posthoc tests show that WI SONRIME is significantly longer than SD ( $p < .001$ ) and ZB ( $p < .05$ ). Women have significantly longer sonorant rime durations than men (on average, 210 ms vs 186 ms),  $p < .001$ . Finally, f-FR sonorant rime durations are significantly longer than f-FALLS (225 vs 196,  $p < .001$ ). nf-FALLS have the shortest sonorant rime (173 averaged over varieties), but are not significantly different from other sentence conditions.

If we look at the sentence conditions separately, we find main effects of DIALECT and GENDER but no interaction for nf-FALLS, f-FALLS and f-FR. Female speakers had significantly longer sonorant rime durations than male speakers in all three sentence conditions. As for DIALECT, post-hoc comparisons showed that the main effect of DIALECT in nf-FALLS was due to the short rime durations in SD compared to the other varieties. In f-FALLS it was due to the difference between ZB and WI, and in f-FR to the difference between SD and WI.

Table 6. *Effect of DIALECT and GENDER on SONRIME in non-final falls, final falls and final fall-rises.*

	DIALECT		GENDER	
<i>nf-FALL</i>	$F(5,104) = 6.35$	$p < .001$	$F(1,104) = 12.20$	$p < .001$
<i>f-FALL</i>	$F(5,76) = 2.61$	$p < .05$	$F(1,76) = 15.29$	$p < .001$
<i>f-FR</i>	$F(5,83) = 2.37$	$p < .05$	$F(1,83) = 9.56$	$p < .01$

Table 7. *Pairwise comparisons for SONRIME between levels of DIALECT, separately for each sentence condition.*

Levels	SD vs. ZB	ZB vs. RO	RO vs. AM	AM vs. GR	GR vs. WI
<i>nf-FALL</i>	**				
<i>f-FALL</i>					
<i>f-FR</i>					
	SD vs. RO	ZB vs. AM	RO vs. GR	AM vs. WI	
<i>nf-FALL</i>	*				
<i>f-FALL</i>					
<i>f-FR</i>					
	SD vs. AM	ZB vs. GR	RO vs. WI		
<i>nf-FALL</i>	**				
<i>f-FALL</i>					
<i>f-FR</i>					
	SD vs. GR	ZB vs. WI			
<i>nf-FALL</i>	**				
<i>f-FALL</i>		*			
<i>f-FR</i>					
	SD vs. WI				
<i>nf-FALL</i>	***				
<i>f-FALL</i>					
<i>f-FR</i>	*				

### 5.3.2 Peak timing

In Hanssen, Peters, and Gussenhoven (ms), which investigated how a number of Dutch dialects respond to IP-final time pressure, we found evidence that speakers time the f0 peak in monosyllabic words such that the same proportion of the word is available for the realization of the falling pitch movement. That is, the peak in shorter words was timed earlier when measured relative to the beginning of the vowel, but no difference in proportional peak timing was found between shorter and longer words.

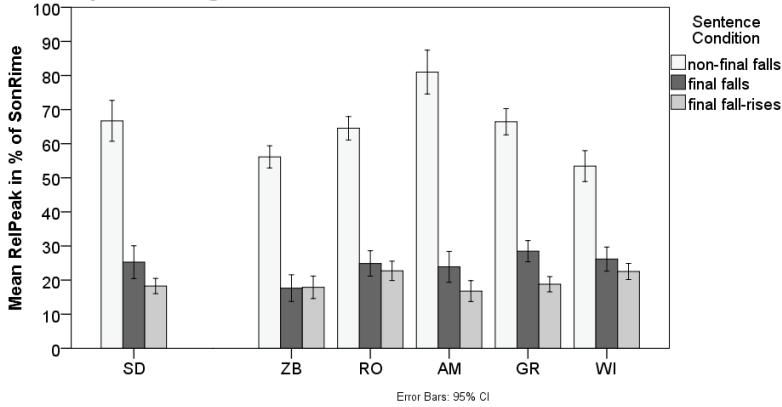
**Relative peak timing in three sentence conditions**

Figure 3. Mean proportional peak timing in non-final falls, final falls and final fall-rises for each variety. Error bars represent the 95% confidence interval.

The bars in Figure 3 suggest that (a) regional differences in proportional peak timing do exist in our data (and are not a mere consequence of differences in segmental duration) and that (b) regional variation in peak timing would appear to interact with contour condition. In non-final falls, peak timing is early in peripheral ZB and WI, intermediate in central RO and GR, also including SD, and late in AM. Closer inspection of the data suggests that in ZB and WI, the peak falls on the nucleus/coda-boundary, whereas in RO, GR and SD it occurs just after the beginning of the coda, and in AM it is timed near the end of the accented syllable. In f-FALLS, the earliest peaks occur in ZB and the latest in GR. In f-FR, RO has the latest peaks, while AM has the earliest.

Table 8 shows main effects of DIALECT and SENTENCE\_CONDITION on proportional peak timing, and interactions between DIALECT x SENTENCE\_CONDITION, GENDER x SENTENCE\_CONDITION and DIALECT x GENDER x SENTENCE\_CONDITION.

Table 8. Effect of DIALECT, GENDER and SENTENCE\_CONDITION on RELPEAK.

<i>Dialect</i>	$F(5,108) = 2.91$	$p < .05$
<i>Gender</i>	$F(1,108) = 1.63$	<i>n.s.</i>
<i>Sentence_condition</i>	$F(2,34) = 408.51$	$p < .001$
<i>Dialect x Sentence_condition</i>	$F(10,974) = 11.32$	$p < .001$
<i>Gender x Sentence_condition</i>	$F(2,982) = 4.19$	$p < .05$
<i>Dialect x Gender x Sentence_condition</i>	$F(10,972) = 1.92$	$p < .05$

Posthoc tests show that the proportional peak in AM is timed significantly later than in ZB ( $p < .01$ ). They also show that all sentence conditions differ

significantly from one another ( $p < .001$  for all levels of the comparison), with estimated mean proportional peak timings of 63% (nf-FALL), 25% (f-FALL) and 20% (f-FR).

Looking at each sentence condition separately, we find a main effect of DIALECT [ $F(5,106) = 9.88, p < .001$ ] and GENDER [ $F(1,106) = 4.44, p < .05$ ], no interaction, in nf-FALLS only. Female speakers time the location of the non-final nuclear peak at around 66% of the sonorant syllable, compared to 60% for male speakers. Posthoc tests for DIALECT show that in nf-FALLS, AM peaks are significantly later than all other varieties, with mean differences ranging from 18 (AM-RO) to 29 per cent (AM-WI).

5.3.3 Scaling of tonal targets

Next, we address differences in scaling of individual pitch targets and overall pitch level. Since GENDER is unbalanced between varieties in our design, and converting from Hz to Semitones will not completely neutralize gender variation, we look at female speakers only for this part of the results. Scaling of the peak (H) and the following low target (L) can be analyzed across sentence conditions, scaling of the final high (H2) only occurs in fall-rises.

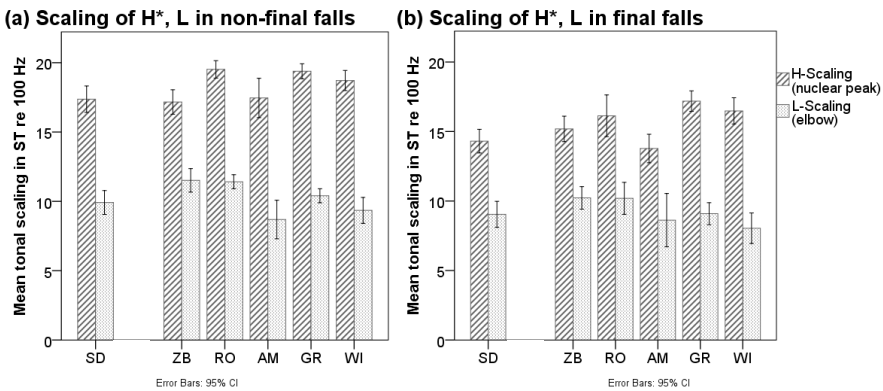


Figure 4. Mean scaling in semitones of H and L in nf-FALLS (left-hand panel) and f-FALLS (right-hand panel) for each variety. Error bars represent the 95% confidence interval.

### Scaling of H\*, L and H% tones in final fall-rises

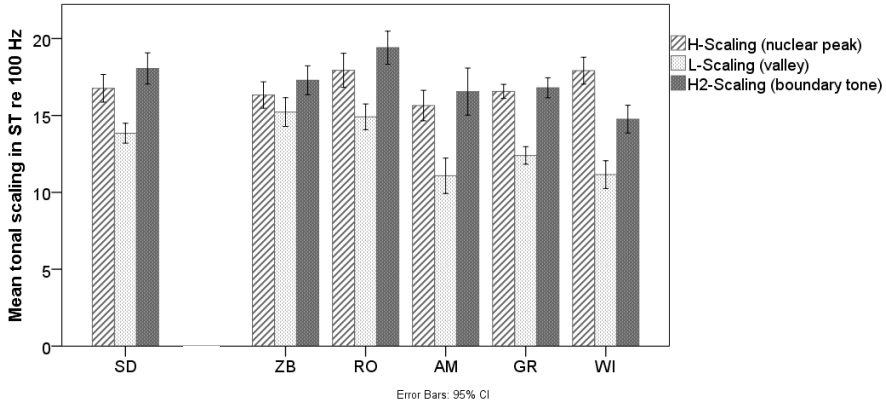


Figure 5. Mean scaling in semitones of H, L and H2 in f-FR for each variety. Error bars represent the 95% confidence interval.

The bar charts in Figs. 4 and 5 do not show a clear regional pattern for scaling of the nuclear peak, although RO, GR and WI tend to be higher than AM in particular. Scaling of L tends to be high in ZB and RO, and lower in AM, GR and WI, with SD in between. Figure 5 further suggests that overall, scaling is low in AM and high in RO. We also see that the final rise of the fall-rise is higher in the central-to-southern varieties than in the east.

Statistical analysis for H-Scaling revealed a main effect of SENTENCE\_CONDITION [ $F(2,25) = 37.78$ ,  $p < .001$ ] and a DIALECT \* SENTENCE\_CONDITION interaction [ $F(10,592) = 5.04$ ,  $p < .001$ ], but no main effect of DIALECT. Posthoc tests show that all sentence conditions differ significantly from one another (all pairs  $p < .001$ ), with highest peaks in nf-FALLS (18.3 ST) followed by f-FR (16.7 ST) and f-FALLS (15.9 ST). Although peaks are always highest in non-final falls, they are not always lowest in final falls, which may have caused the interaction.

As for scaling of the low target (L), we found a main effect of DIALECT [ $F(5,64) = 2.97$ ,  $p < .05$ ], of SENTENCE\_CONDITION [ $F(2,13) = 152.40$ ,  $p < .001$ ] and a DIALECT \* SENTENCE\_CONDITION interaction [ $F(10,592) = 4.83$ ,  $p < .001$ ]. Post-hoc tests revealed no significant differences in L-SCALING between any of the dialects, but showed that L in fall-rises was significantly higher (13.1 ST) than both non-final (10.3 ST) and final falls (9.6 ST) at  $p < .001$ .

Separate analyses for final and non-final falls showed that H-SCALING was significantly affected by DIALECT in f-FALLS [ $F(5,48) = 2.71$ ,  $p < .05$ ], although none of the varieties differed significantly in post-hoc tests.

In final fall-rises, DIALECT did not significantly affect H-SCALING, but as table 9 shows, we did find a main effect of DIALECT on L-SCALING and H2-SCALING.

Table 9. *Effects of DIALECT on the scaling of the nuclear peak, the elbow and the final high target in final fall-rises.*

H-SCALING	$F(5, 59) = 1.33$	n.s.
L-SCALING	$F(5, 59) = 5.17$	$p < .001$
H-SCALING	$F(5, 59) = 3.37$	$p < .01$

Bonferroni post-hoc comparisons showed that L was significantly higher in ZB compared to AM ( $p < .05$ ) and WI ( $p < .01$ ), and significantly higher in RO compared to WI ( $p < .05$ ). Scaling of the second high target H2 was significantly higher in SD compared to WI ( $p < .05$ ) and RO compared to WI ( $p < .01$ ).

Even though our analyses do not reveal a systematic effect of DIALECT on the nuclear peak, inspection of Figs. 4 and 5 suggests that differences in excursion size exist between varieties. We generally see wide falls in the north-east, and small excursions in the south-west. Section 5.3.4 looks into such differences in contour shape.

### 5.3.4 Contour shape: f0 duration, excursion and slope

In this part of the results, we compare the shape of nuclear contours between varieties by looking at f0 duration, f0 excursion and slope of the pitch movements. Note that these variables are not comparable between non-final and final falls on the one hand, and fall-rises on the other. For falls, the variables apply to the f0 stretch between the peak (H\*) and the subsequent elbow (L), disregarding the level or slowly falling pitch after the elbow. For fall-rises, the variables are measured separately for the falling (H\* to L) and the rising (L to H%) part of the pitch contour. We will therefor present the results for contour shape separately for falls (section 5.3.4.1) and fall-rises (5.3.4.2). GENDER returns as an independent variable, since differences in excursion are largely neutralized by measuring f0 in semitones.

#### 5.3.4.1 Non-final and final falls

As the bar charts in Figure 6 clearly show, both f0 duration, excursion and slope vary with dialect and position. Generally, the eastern varieties have wide and steep falling movements. We also see that for all varieties, f0 duration is longer, f0 excursion larger, and by consequence f0 slope is less steep in non-final falls, compared to final falls. Indeed, we find significant effects of POSITION for all three variables, along with significant effects of DIALECT and



DIALECT\*POSITION interactions. The difference between the two positions is larger in some varieties than in others (compare RO and WI, for example). Amsterdam stands out in this respect, with particularly long  $f_0$  durations and shallow slopes in nf-FALLS. We will come back to this result in the discussion. Below, we look at the effect of DIALECT on contour shape separately for non-final and final falls.

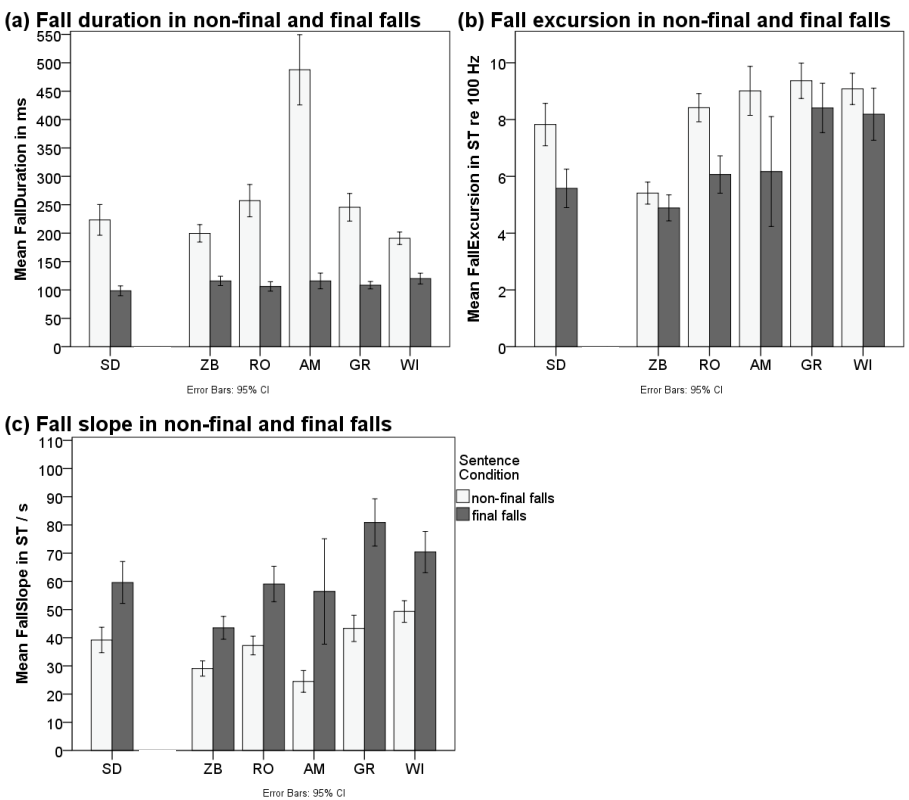


Figure 6. Mean  $f_0$  duration in ms (panel a),  $f_0$  excursion in ST (panel b) and  $f_0$  slope in ST/s (panel c) for non-final and final falls, broken down by dialect. Error bars represent the 95% confidence interval.

In nf-FALLS, we found a significant main effect of DIALECT on FALLDURATION, FALLEXCURSION and FALLSLOPE. GENDER did not significantly affect any of the dependent variables. The results are summarized in table 10. Bonferroni posthoc comparisons for FALLDURATION revealed that AM had significantly longer durations than all other varieties at the  $p < .001$  level. None of the other varieties differed significantly from one another. For FALLEXCURSION, posthoc tests show that ZB excursion is significantly smaller

than all other varieties (ZB-SD and ZB-RO  $p < .05$ ; ZB-AM and ZB-GR  $p < .001$ ; ZB-WI  $p < .01$ ). Finally, posthoc tests for FALLSLOPE show that falls are significantly less steep in ZB compared to GR ( $p < .001$ ) and WI ( $p < .05$ ), and less steep in AM compared to all varieties except ZB (AM-SD  $p < .01$ ; AM-RO  $p < .05$ ; AM-GR and AM-WI  $p < .001$ ).

Table 10. *Effects of DIALECT on FALLDURATION, FALLEXCURSION and FALLSLOPE in non-final falls and final falls.*

FALLDURATION		
<i>nf-FALL</i>	$F(5,107) = 24.10$	$p < .001$
<i>f-FALL</i>	$F(5,76) = 3.18$	$p < .05$
FALLEXCURSION		
<i>nf-FALL</i>	$F(5,103) = 6.51$	$p < .001$
<i>f-FALL</i>	$F(5,76) = 4.52$	$p < .001$
FALLSLOPE		
<i>nf-FALL</i>	$F(5,105) = 8.45$	$p < .001$
<i>f-FALL</i>	$F(5,77) = 5.19$	$p < .001$

Continuing with f-FALLS, table 10 shows main effects of DIALECT for all three variables. Additionally, we found an effect of GENDER on FALLSLOPE in final falls [ $F(1,77) = 4.18$ ,  $p < .05$ ]. Bonferroni comparisons for FALLDURATION revealed one significantly different dialect pair: SD-WI ( $p < .05$ ). For both FALLEXCURSION and FALLSLOPE, we found that GR had significantly larger and steeper falls than SD ( $p < .05$  for both variables) and ZB ( $p < .01$  and  $p < .001$ , respectively).

### 5.3.4.2 Fall-rises

Regional differences in the shape of the fall-rise may be due to differences in the shape of the falling movement (H\* to L), of the final rise (L to H%), or both. We therefore measured pitch movement duration, excursion and slope separately for the falling movement (FALLDURATION, FALLEXCURSION and FALLSLOPE) and the final rise (RISEDURATION, RISEXCURSION and RISESLOPE). The bar charts in Figure 7 show that there are rather large differences in the shape of the fall-rise, with ZB, GR and WI differing most from the other varieties, each in their own way. For the falling movement, WI has the longest duration, largest excursion and steepest slope. ZB, on the other hand, has the shortest duration, smallest excursion and shallowest slope for both the falling and the rising movement. In GR, the difference between the falling and rising part is small in terms of duration, excursion and slope. These examples illustrate that regional patterns vary for the falling movement, the rising movement and the relation between the two.

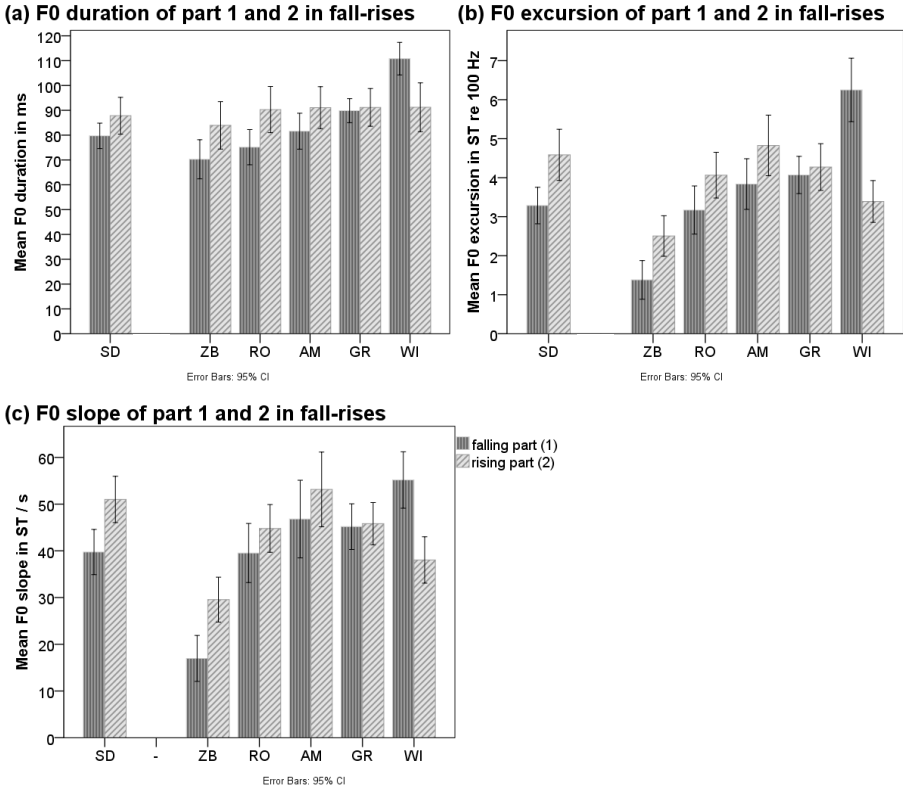


Figure 7. Mean f0 duration in ms (top left panel), f0 excursion in ST (top right panel) and slope in ST/s (bottom panel) of the falling (FR1) and rising (FR2) movements of final fall-rises. Error bars represent 95% of the confidence interval.

We found a main effect of DIALECT on FALLDURATION [ $F(5,83) = 2.66, p < .05$ ] and RISESLOPE [ $F(5,83) = 2.94, p < .05$ ]. We also found a main effect of GENDER on RISEDURATION [ $F(1,85) = 4.69, p < .05$ ] and a GENDER\*DIALECT interaction [ $F(1,84) = 2.42, p < .05$ ] for FALLEXCUSION.

Inspection of the data showed that female speakers had significantly longer final rise durations than male speakers (88 vs 78 ms). Posthoc comparisons for FALLDURATION showed that f0 duration of the falling movement was significantly longer in WI compared to ZB and RO (both at  $p < .05$ ). For RISESLOPE, no dialect pairs were significantly different in posthoc tests. The interaction between DIALECT and GENDER for FALLEXCUSION may have been caused by the fact that female speakers had a larger excursion in some (WI, GR) but not in other dialects (SD, ZB).

Figure 7 also suggests that regional differences exist in the way the fall and rise of H\*L H% relate to each other. In SD, ZB, RO and AM, the falling movement

is shorter, smaller and shallower than the rising movement. Moving eastwards, the two movements are almost equal in duration, excursion and slope in GR. Finally, in WI, the pattern is the opposite of the south-western and central varieties, with the falling movement being longer, larger and steeper than the final rise.

Figure 8, which shows the relation, or *ratio*, between the falling and the rising movement in each dialect, further illustrates these regional differences in shape of the fall-rise. We computed three variables that reflect the ratio of the falling movement compared to the final rising movement in terms of (1) duration (RATIOFRDUR), (2) excursion (RATIOFREXC) and (3) slope (RATIOFRSLOPE), by dividing the value of the falling movement by that of the rising movement for each speaker. A ratio of  $< 1$  means that the falling part is shorter (or smaller, or shallower) than the rising part. For example, a RATIOFREXC of 0.5 (as in ZB below) means that the falling part is half the size of the rising part.

Statistical analyses revealed a significant main effect of DIALECT on RATIOFRDUR [ $F(5,86) = 2.67, p < .05$ ] and RATIOFREXC [ $F(5,83) = 2.84, p < .05$ ]. Posthoc tests for the duration ratio did not reveal significantly different dialect pairs. For the excursion ratio, they showed that WI is significantly different from SD ( $p < .05$ ).

**Fall / Rise ratio for f0 duration, excursion and slope in fall-rises**

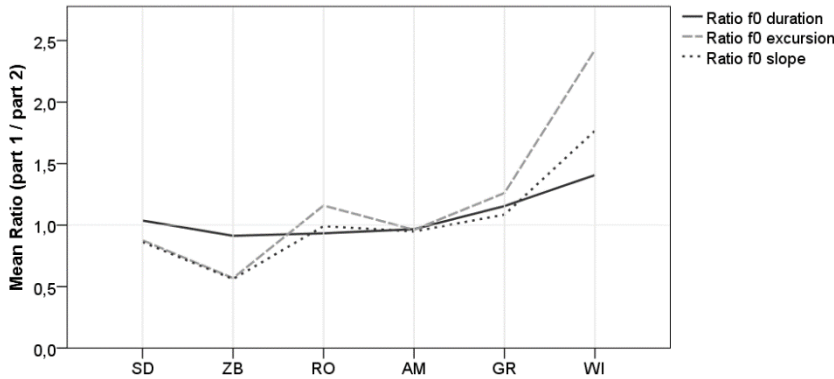


Figure 8. Duration ratio, excursion ratio and slope ratio between the falling and rising movement of final fall-rises.

## 5.4 Summary and discussion

Earlier investigations of regional differences in the realization of intonation have mainly reported variation in tonal timing. Our more comprehensive study included other variables, such as segmental and contour duration, scaling and

contour shape. We will sum up the main results, paying attention to the effect of gender (5.4.1), sentence condition (5.4.2) and dialect (5.4.3). We will compare our findings to the results reported in Peters et al. (2014) to see if we replicate their finding of a geographical cline in the realization of non-final and IP-final falling pitch accents, and IP-final falling-rising pitch accents.

### 5.4.1 Effects of gender

Apart from the inherent effect on scaling, the most systematic effect of gender was found for segmental duration. Rime durations were longer for women across sentence conditions. Women also timed their peaks later than men, most notably in non-final falls, which may be an effect of the increased availability of sonorant material. In final falls, male speakers showed a steeper falling slope, although closer inspection of the data showed that this was not consistent across dialects. Finally, we found that women produced a longer, and in some dialects also larger, final rise in H\*L H% nuclear accents. These features may be explained as an enhanced use of the high-pitched end of Ohala's (1983) Frequency Code (Gussenhoven 2016).

### 5.4.2 Effects of sentence condition

The three sentence conditions varied systematically with respect to sonorant rime duration and timing of the nuclear peak. Rime durations were shorter, and peak timing later, in non-final falls compared to final falls, and again in final falls compared to final fall-rises. Peaks were also scaled higher in non-final contours, compared to final contours. All these effects can be interpreted as responses to time pressure. Speakers can increase the duration of segmental material, and retract and lower tonal targets to create more space for realization of contours in case of an upcoming IP-final boundary or in case of more complex intonation contours (e.g., Wightman et al. 1992, Steele 1986, Prieto et al. 1995, Grabe 1998, also see Hanssen, Peters, and Gussenhoven, ms, for a comparison of such time pressure effects in varieties of Dutch). Final and non-final falls could additionally be compared in terms of shape. Final falls were realized with shorter  $f_0$  durations, shorter excursions and hence steeper falling slopes, all of which can also be attributed to the lack of space to realize the contours.

The realization of nuclear contour types was not affected uniformly across the dialects. First, the difference in sonorant rime duration between the two falling contours is much smaller in Zuid-Beveland than in other varieties, particularly Winschoten. Secondly, the difference between non-final and final proportional peak timing in falls is much smaller in Winschoten than in Standard Dutch, and particularly Amsterdam (see section 5.4.3). In fall-rises, the difference in timing between final falls and fall-rises in Zuid-Beveland is particularly small compared

to Grou. Finally, while final falls had a steeper falling slope in all varieties, differences could be observed in the sources that governed the slope. For example, excursion sizes varied much less between non-final and final falls in Grou and Winschoten, while fall duration was reduced. In Standard Dutch and Rotterdam, steeper slopes were also caused by shorter durations, while excursions were reduced at the same time.

### 5.4.3 Effects of dialect

#### *Duration*

Sonorant rime durations gradually increased from the south-west (ZB) to the north-east (GR, WI), showing a weak geographical component. ZB generally had the shortest durations, and WI the longest, matching the first half of the inverted U-shape reported in Peters et al. (2014). The short segmental durations for ZB and long ones in GR are in agreement with that study, which investigate the effects of focus condition on the realization of non-final declarative falls in varieties of Dutch (but which did not look at Standard Dutch or the Winschoten variety).

Recall from section 5.2.3 that speakers from WI often pronounced the target words, e.g. ‘Manderen’ as a disyllabic word [mandə:n] instead of [mandərə]. This reduction may partly explain the long rime durations, since the fewer unstressed syllables in a word, the longer its stressed syllable (Nooteboom 1972; Rietveld, Kerkhof, and Gussenhoven 2004), meaning that WI speakers pronounce a longer nuclear syllable than other varieties because this syllable is followed by one as opposed to two unstressed syllables. However, in the case of final falls and final fall-rises, with identically pronounced target words in all varieties, segmental duration was always longest in WI. We conclude from this that, apart from any metrical effect on syllable duration, speakers of WI have longer segmental durations than other varieties, giving them a slower overall speech rate. This finding is reminiscent of work by Verhoeven, De Pauw and Kloots (2004<sup>8</sup>), who reported that speech tempo was slowest for speakers from the northern and south-eastern peripheries of the Dutch language area. We note that our data from south-western Zeeland do not conform to this trend.

#### *Peak timing*

DIALECT had an effect on proportional peak timing, although it was less general than its effect on sonorant rime duration. Details differed with contour condition. In non-final falls, proportional peak timing tended to be earliest in ZB and WI

---

<sup>8</sup> Verhoeven et al.’s work was criticized for methodological errors in Quené (2008), who nevertheless reached the same conclusion regarding differences in speech tempo between speakers from Flanders and the Netherlands, and between male and female speakers.

and latest in AM, a geographical pattern that closely resembles the results in Peters et al. (2014). The other contour classes showed a different pattern, although ZB always belonged to the early-peak group. In WI, peaks were late in the two final conditions but relatively early in non-final falls. Since the duration of the final accented syllable rime was particularly long in WI, the late timing of the final peaks can be attributed to the generous availability of sonorant segments.

Its late-peak accents make AM stand out particularly from the other varieties. On closer inspection of the data, we observed that speakers produced a combination of normal-peak and late-peak falling accents, with or without a final boundary tone (i.e., H\*L L%, H\*L 0% or L\*HL 0%) for non-final declarative falls. Due to large between- and within-speaker variation, we could often not confidently label the pitch accent as either a late or a regular fall. In Hanssen, Gussenhoven, and Peters (ms), the AM late-peak accents are studied in more detail.

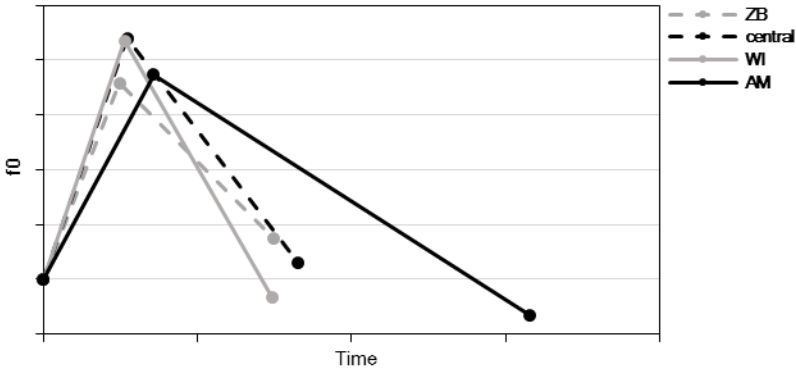
### *Scaling*

Looking across sentence conditions, DIALECT did not systematically affect the scaling of nuclear peaks or overall pitch level, as was found between the standard varieties of Dutch in the Netherlands and Belgium (van Bezooijen 1993). The largest effect of DIALECT on scaling was found in the fall-rises, where the valley between the two high tones was realized at much higher f<sub>0</sub> in ZB and RO than in AM and WI. The final high tone was scaled highest in WI. Differences in the depth of the valley and height of the final high tone have consequences for the excursion sizes of the falling and rising movements of the fall-rise, as Figure 7 suggests.

### *Contour shape*

Starting with the effect of DIALECT on the shape of nuclear falling melodies, we have seen that the fall (from nuclear peak to subsequent elbow) was particularly long and shallow in non-final position in Amsterdam. This follows directly from the presence of late-peak, slowly falling nuclear accents in that variety. Speakers use the late-peak accent to a much lesser extent if the nuclear accent is IP-final, which puts pressure on its realization. Furthermore, we see that the southwestern ZB falls tend to be small and shallow, in sharp contrast with the large and steep falls in the north-east. Comparing our non-final data with the results for contour shape in Peters et al. (2014), we also see an inverted U-shape for f<sub>0</sub> excursion. The U-shape for f<sub>0</sub> slope that was found in Peters et al. (2014) is not replicated in our findings, because fall slope for ZB is shallow in our results, whereas it is steep in theirs. Apart from that difference, the dialects show similar behavior in both studies. Figure 9 illustrates the shape of non-final (panel a) and final falls (panel b).

(a) Illustration of non-final fall shape



(b) Illustration of final fall shape

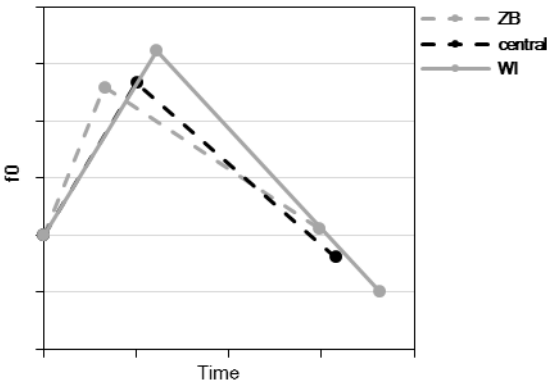


Figure 9. Schematic representation of non-final fall shapes (panel a) in peripheral (Zuid-Beveland and Winschoten) and central varieties, and in Amsterdam, and of final fall shapes (panel b) in peripheral (ZB and WI) and central varieties (including AM).

We continue with the fall-rise, which was rather similar in shape in SD, RO, AM and GR (see Figure 7 in section 5.3.4.2). We call this the ‘central’ realization of the fall-rise. ZB and WI deviated from this ‘central’ pattern in their own ways. Speakers of WI realized the fall-rise with longer, larger and steeper falling movements and shorter and shallower final rises than speakers of the other varieties. As a result, the ratio between the falling and rising movements in WI differed substantially from that in the other varieties. In the other varieties, the shape of the falling and rising movements were comparable (GR), or else the fall was short, small and shallow compared to the final rise. The stylized contours in Figure 10 show the distinct shape of the fall-rise in WI. In striking contrast, the excursion and slope of both the falling and the rising movements in ZB were



considerably smaller and shallower than in the central varieties. As shown in Figure 10, the shape of the ZB fall-rise is characterized by a shallow dip between the two high peaks. In a paper looking at dialect-specific nuclear contours included the ZB variant of the fall-rise, it is shown that speakers of ZB often do not produce such a dip at all (Hanssen, Gussenhoven, and Peters ms). These extremely shallow realizations by speakers of Zuid-Beveland represent a context-specific response to time pressure that is absent in the other varieties.

**Illustration of final fall-rise shape**

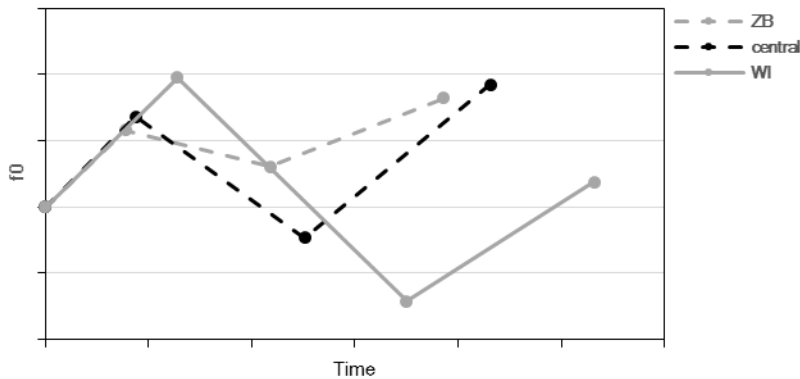


Figure 10. Schematic representation of fall-rise types in peripheral (Zuid-Beveland and Winschoten) and central varieties.

Significantly, the most extreme realizational variation could be observed in the two geographically most extreme dialects. Zuid-Beveland and Winschoten, which are both located in the periphery of the Netherlands, differed most from each other as well as from the other varieties. In fact, if we look at the significantly different dialect pairs, 67% of them (29 out of 43) involve ZB, WI or both. In only 14 out of 43 cases are dialects other than ZB and WI involved in the comparison. Thus, Zuid-Beveland and Winschoten represent two extreme ends of a scale with respect to segmental duration, and excursion and shape of the pitch movements, with Standard Dutch, Rotterdam, Amsterdam and Grou generally in between. Interestingly, the variety of Weener Low Saxon in Germany can often be placed geographically and linguistically after Winschoten, logically extending the inverted U-shape across the border, where it prematurely ended in our study.

Except for variables related to the AM late fall, dependent variable means in SD, RO and AM often resembled one another. This reflects the fact that the standard variety has its roots in the central western varieties of the Netherlands. In fact, for many variables, we could observe a tendency for a gradual shift in the mean

from the southeastern ZB via the central varieties of RO, AM and SD, and on to the northeastern varieties of GR and WI. Such geographical clines had been reported for the first time in Peters et al. (2014, 2015) for Dutch, Frisian, and Low and High Saxon intonation. We have presented additional evidence to support the existence of such a geographical cline, which are commonplace in segmental sociolinguistic research, *modulo* social and natural boundaries (Britain 2013) but new in the field of intonation.

## References

- Atterer, M. and Ladd, D.R. (2004). On the phonetics and phonology of “segmental anchoring” of F0: evidence from German. *Journal of Phonetics* 32, 177–197.
- Arvaniti, A. and Garding, G. (2007). Dialectal variation in the rising accents of American English. In: Hualde, J. and Cole, J. (Eds.), *Papers in Laboratory Phonology* 9. Berlin: Mouton de Gruyter, 547–576.
- Avesani, C. and Vayra, M. (2003). Broad, narrow and contrastive focus in Florentine Italian. In: *Proceedings of the 15th International Conference of Phonetic Sciences (ICPhS), Barcelona*, 1803–1806.
- Boersma, P. and Weenink, D. (2008). Praat: doing phonetics by computer (Version 5.0.25) [computer program]. Retrieved 31 May 2008 from <http://www.praat.org/>.
- Britain, David (2013). Space, Diffusion and Mobility. In: Chambers, J.K. and Schilling, N. (Eds.), *The Handbook of Language Variation and Change* (2nd edition). Hoboken, New Jersey: Wiley-Blackwell, 471–500.
- Daan, J. (1938). *Dialect and pitch pattern of the sentence*. Proceedings of the 3rd international congress of phonetic sciences, Ghent, 473–480.
- Dalton, M. and Ní Chasaide, A. (2005). Tonal alignment in Irish dialects. *Language and Speech* 48, 441–464.
- Dalton, M. and Ní Chasaide, A. (2007). Nuclear accents in four Irish (Gaelic) dialects. In: *Proceedings of the 16th International Conference of Phonetic Sciences (ICPhS), Saarbrücken*, 965–968.
- Del Giudice, A., Shosted, R., Davidson, K., Salihie, M., and Arvaniti, A. (2007). Comparing methods for locating pitch “elbows”. In: *Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS), Saarbrücken*, 1117–1120.
- Gilles, P. (2005). *Regionale Prosodie im Deutschen: Variabilität in der Intonation von Abschluss und Weiterweisung*. Berlin: Walter de Gruyter.
- Grabe, E. (1998). Pitch accent realization in English and German. *Journal of Phonetics* 26, 129–144.
- Grabe, E. (2004). Intonational variation in urban dialects of English spoken in the British Isles. In: Gilles, P. and Peters, J. (Eds.), *Regional variation in intonation*. Tübingen: Niemeyer, 9–31.
- Grabe, E. and Post, B. (2002). Intonational variation in the British Isles. In: Bel, B. and Marlien, I. (Eds.), *Proceedings of Speech Prosody 2002*, Aix-en-Provence, 343–346.
- Grabe, E., Post, B., Nolan, F., and Farrar, K. (2000). Pitch accent realization in four varieties of British English. *Journal of Phonetics* 28, 161–185.
- Gussenhoven, C. and van der Vliet, P. (1999). The phonology of tone and intonation in the Dutch dialect of Venlo. *Journal of Linguistics* 35, 99–135.
- Gussenhoven, C. (2016). Foundations of intonational meaning: Anatomical and physiological factors. *Topics in Cognitive Science* 8, 425–434.
- Hanssen, J., Gussenhoven, C., and Peters, J. (ms). *Melodic aspects of dialect-accented Dutch*.
- Hanssen, J., Peters, J., and Gussenhoven, C. (ms). *Regional variation in phonetic responses to time pressure in Dutch IP-final nuclear contours*.
- Hanssen, J., Peters, J., and Gussenhoven, C. (2016). Phonetic effects of focus in five varieties of Dutch. *Proceedings of Speech Prosody 2016, Boston*, 736–740.

- Hirst, D.J. (2005). Form and function in the representation of speech prosody. *Speech Communication* 46, 334-347.
- Kalaldehy, R., Dorn, A., and Ní Chasaide, A. (2009). Tonal alignment in three varieties of Hiberno-English. In: *Proceedings of Interspeech 2009, Brighton*, 2443-2446.
- Kügler, F. (2007). *The intonational phonology of Swabian and Upper Saxon*. PhD thesis. Tübingen: Max Niemeyer Verlag.
- Ladd, D.R. (2008). *Intonation phonology*. 2<sup>nd</sup> edition. Cambridge: Cambridge University Press.
- Ladd, D.R., Schepman, A., White, L., Quarmby, L.M., and Stackhouse, R. (2009). Structural and dialectal effects of pitch peak alignment in two varieties of British English. *Journal of Phonetics*, 37, 145-161.
- Mücke, D., Grice, M., Becker, J., and Hermes, A. (2009). Sources of variation in tonal alignment: Evidence from acoustic and kinematic data. *Journal of Phonetics* 37, 321-338.
- Nooteboom, S.G. (1972). *Production and perception of vowel duration. A study of durational properties of vowels in Dutch*. University of Utrecht PhD thesis.
- O'Reilly, M., Dorn, A., and Ní Chasaide, A. (2010). Focus in Donegal Irish (Gaelic) and Donegal English bilinguals. In: *Proceedings of Speech Prosody 2010, Chicago*.
- Peters, J. (1999). The timing of nuclear high accents in German dialects. In: *Proceedings of the 14th International Congress of Phonetic Sciences (ICPhS)*, San Francisco, 1877-1880.
- Peters, J. (2006). *Intonation deutscher Regionalsprachen*. Berlin: de Gruyter.
- Peters, J., Hanssen, J., and Gussenhoven, C. (2014). The phonetic realization of focus in West Frisian, Low Saxon, High German, and three varieties of Dutch. *Journal of Phonetics* 46, 185-209.
- Peters, J., Hanssen, J., and Gussenhoven, C. (2015). The timing of nuclear falls: Evidence from Dutch, West Frisian, Dutch Low Saxon, German Low Saxon, and High German. *Laboratory Phonology* 6, 1-52.
- Peters, J. and Gussenhoven, C. (ms). *Intonational variation in the Netherlands and beyond. A study of Zeelandic, Hollandic Dutch, West Frisian, Dutch Low Saxon, German Low Saxon, and High German*.
- Petrone, C. and D'Imperio, M. (2009). Is tonal alignment interpretation independent of methodology? In: *Proceedings of Interspeech 2009, Brighton*, 2459-2462.
- Prieto, P., D'Imperio, M., and Gili Fivela, B. (2005). Pitch accent alignment in Romance: primary and secondary associations with metrical structure. *Language and Speech* 48, 359-396.
- Prieto, P., van Santen, J., and Hirschberg, J. (1995). Tonal alignment patterns in Spanish. *Journal of Phonetics* 23, 429-451.
- Quené, H. (2008). Andante of allegro? Verschillen in spreektempo tussen Vlamingen en Nederlanders. *Onze Taal* 77, 179-181.
- Rietveld, T., Kerkhoff, J., and Gussenhoven, C. (2004). Word prosodic structure and vowel duration in Dutch. *Journal of Phonetics* 32, 349-371.
- Smakman, D. (2006). *Standard Dutch in the Netherlands. A sociolinguistic and phonetic description*. PhD thesis. Utrecht: LOT publications.
- Steele, S. (1986). Nuclear accent *f<sub>0</sub>* peak location: Effect of rate, vowel, and number of syllables. *Journal of the Acoustical Society of America* 80 (Suppl. 1), s51.

- Turk, A., Nakai, S., and Sugahara, M. (2006). Acoustic segment durations in prosodic research: a practical guide. In: Sudhoff, S., Lenertová, D., Meyer, R., Pappert, S., Augurzky, P., Mleinek, I., Richter, N., and Schliesser, J. (Eds), *Methods in empirical prosody research*. Berlin, New York: De Gruyter (= Language, Context, and Cognition, 3), 1-28.
- Ulbrich, C. (2005). *Phonetische Untersuchungen zur Prosodie der Standardvarietäten des Deutschen in der Bundesrepublik Deutschland, in der Schweiz und in Österreich*. PhD thesis. Frankfurt am Main: Lang.
- van Bezooijen, R. (1993). Verschillen in toonhoogte: Natuur of cultuur? *Gramma/TTT* 2, 165-179.
- van Es, G.A. (1935). Syntactische functies der intonatie in de volkstaal onzer noordelijke provinciën. *Handelingen van het Zestiende Nederlandsche Philologen-Congres*, 39-42.
- van Leyden, K. (2004). *Prosodic characteristics of Orkney and Shetland dialects: An experimental approach*. PhD thesis Leiden University.
- Verhoeven, J., de Pauw, G., and Kloots, H. (2004). Speech rate in a pluricentric language situation: a comparison between Dutch in Belgium and the Netherlands. *Language and Speech* 47, 299-310.
- Weijnen, A. (1966). *Nederlandse dialectkunde*. Assen: Van Gorcum.
- Wightman, C.W., Shattuck-Hufnagel, S., Ostendorf, M., and Price, P.J. (1992). Segmental durations in the vicinity of the prosodic phrase. *Journal of the Acoustical Society of America* 91, 1707-1717.



# NON-STANDARD MELODIES AND MELODY PREFERENCES IN DIALECT-ACCENTED DUTCH<sup>1</sup>

---

## Chapter 6

### Abstract

In this study, dialect-specific preferences for IP-final and IP-medial nuclear contours in different pragmatic contexts are examined. Speakers from five locations along the Dutch coastal line, as well as a group of Standard Dutch speakers, participated in a reading task. Local varieties included Zeelandic Dutch, South Hollandic and North Hollandic, West-Frisian, and Low Saxon. We found that all nuclear contours observed in the six varieties have been reported to exist in Standard Dutch, and therefore conclude that they share a tonal grammar. However, our results also show that varieties vary semantically, as evidenced by the variation in the distribution of nuclear contours over sentence types, and in the frequency with which the preferred melodies are used. Particularly, Zeelandic Dutch preferred H\*L L% for interrogatives, which were mostly produced with rising or falling-rising intonation in the other varieties. And whereas (late) half-completed falls could be observed on declaratives in all varieties, its use was particularly frequent in the Hollandic varieties. Finally, this paper provides a detailed account of three dialect-specific contour realizations, which are the *interrogative fall* and the *rise-rise* from Zeelandic Dutch, and the *(delayed) half-completed fall* from Rotterdam and Amsterdam Dutch. We conclude that both distributional differences and dialect-specific phonetic realization can contribute to a dialect's 'signature'.

### 6.1 Introduction

It has long been assumed that Dutch dialects have characteristic melodies, and that intonation may be the most important cue to a speaker's (linguistic and geographical) origin (e.g., van Es 1932:93). To quote Daan (1938:473):

"[...] it is possible to recognize a dialect by other characteristics than only by the words and sounds, [but] it is very difficult to say which these characteristics are. No doubt the musical accent plays an important part in this matter."

---

<sup>1</sup> This chapter is a slightly revised version of: Hanssen, J., Gussenhoven C., and Peters, J. (ms). "Non-standard melodies and melody preferences in dialect-accented Dutch". Submitted to *Language and Speech*.

Daan (1938) claimed that speakers of Zeeland, Noord-Holland and Friesland have ‘excessive musical accent’, whereas those of West-Brabant, Zuid-Holland, Utrecht and the Northeast of the Netherlands have ‘normal pitch patterns’. She considered the dialect of Drenthe monotonous, that slowness is probably typical of the North Hollandic dialects, and that Frisian is characterized by big intervals and rising at the utterance end. Similarly, Weijnen (1966) and, more recently, Wortel (2002) and van Oostendorp (2002), make impressionistic reference to dialectal variation in intonation. Limburgian, North Hollandic, and the Rotterdam urban dialect are described as having ‘lilting’ intonation. Van Oostendorp (2002:37) attributes this impression to the fact that we usually compare dialects to Standard Dutch, which he describes as relatively flat. He regrets that his comments are impressionistic, adding that “unfortunately no scientific research has been conducted on the differences in intonation of the different city dialects”. This comment applies to all regional varieties of the Netherlands, with the exception of the tonal dialects spoken in Limburg.

Some evidence in support of van Es’ claim that a speaker’s origins is traceable in his intonation is provided by Gooskens (1997). A number of perception experiments suggested that both exogenous listeners (Standard Dutch speakers) and endogenous listeners (dialect speakers) used prosodic information to distinguish their own variety from other varieties. For varieties other than their own, speakers of Standard Dutch could not identify language varieties on the basis of prosodic information, although prosodic information did improve language identification if it was combined with verbal information.

We have two aims in this contribution. In terms of Ladd’s taxonomy of intonational variation (1996:119/2008:116), the first aim focuses on semantic/systemic variation and the second on realizational variation. We first present dialect-specific preferences in six varieties of Dutch for nuclear contours in different pragmatic contexts. Grabe and Post (2002) and Grabe (2004) found that varieties of British English differed in the frequency with which they used specific nuclear melodies. While the varieties shared the same set of pitch accents and boundary tones, they appeared to be characterized by what might be referred to as pitch accent/nuclear tune frequency profiles (NTFP). We decided to investigate the NTFPs and dialect-specific intonation contours for IP-final syllables in different pragmatic contexts. Unlike non-final accented syllables in the IP, IP-final syllables comprise the entire pitch contour corresponding to the pitch accent plus the IP-final boundary tone and as such represent the most salient location for intonational melodies. The second aim is to present intonation contours which were found in one or two varieties only. Because the Amsterdam and Rotterdam dialects yielded falling contours which differed from the equivalent standard Dutch contours in both IP-non-final and IP-final accented syllables, we include data for statements with IP-medial accented syllables in this report. In view of the long history of comments about



intonational differences in Dutch, it is noteworthy that, as far as we are aware, this is the first time that dialect-specific melodies have been identified and described.

## 6.2 Procedure

Our data were obtained with the help of a reading task based on a corpus of short dialogues. Data for the local varieties were obtained in the localities concerned. This section describes the materials, the subjects and the recording procedure.

### 6.2.1 Materials

For each position in the IP, we selected 12 carrier sentences. Those for the IP-final pitch accents were equally divided over three pragmatic conditions: declarative sentences, yes/no-questions and rhetorical questions. These data will be used for answering the question whether there are systemic/semantic differences among the dialects we investigated. Section 6.3 describes the phonological contours speakers used and what their frequency distributions were in each of these categories. The 12 carrier sentences for the IP-medial pitch accent were all declarative sentences. While these data will be considered in section 6.3, they mainly served to provide data for investigating the variation in falling contours in Amsterdam and Rotterdam which are reported in section 6.4, together with data for rising contours from Zuid-Beveland. Section 6.5 provides a conclusion.

All 24 carrier sentences (labeled ‘B’) were preceded by a context sentence (‘A’) with which they formed a mini-dialogue, as illustrated for the four data sets in (1). In the 12 non-final statements, each set of three carrier sentences contained the fictitious place names *Momberen*, *Memberen*, *Manderen* and *Munderen*, which had the metrical pattern *snm*, in which the segmental structure of the accentable first syllable was Nasal-V-Nasal, followed by a voiced plosive onset consonant. **These target words were followed by two verbs with the pattern *sw*.** In the 12 carrier sentences with accentable IP-final positions, four fictitious monosyllabic proper names, *lof*, *loof*, *lom*, *loom*, were used as target words in each of the three pragmatic conditions. These varied in the rime only, where short [ɔ] and long [o:] combined with voiceless [f] and sonorant [m]. Accented syllables are underlined in (1). We translated the sentences into the local language for speakers from West Frisian and Low Saxon, who have their own standardized spelling system. We used the Dutch version of the sentences for the other speakers. (For more information on our speakers see section 6.2.2.) An overview of the sentences in all language versions is given in the Appendix.

(1) **non-final statements (12)**

- A. Wat zijn de plannen voor morgen?  
*What are the plans for tomorrow?*
- B. Ik zou wel naar Momberen willen fietsen.  
*I'd like to cycle to Momberen.*

**final statements (4)**

- A. Met wie gaat je baas morgen trouwen?  
*Who is your boss marrying tomorrow?*
- B. Hij trouwt met mevrouw de Lom.  
*He's marrying Mrs. de Lom.*

**final yes/no-questions (4)**

- A. Ik zag net je broer Koen met je buurvrouw langslopen.  
*I just saw your brother Koen walk by, with your neighbor.*
- B. Liep-ie naast mevrouw de Lom? Wat raar, die kennen elkaar toch niet?  
*Did he walk next to Mrs. de Lom? How strange, they don't know each other, do they?*

**final rhetorical questions (4)**

- A. Pepijn de Heer komt straks ook naar 't feest.  
*Pepijn de Heer is also coming to the party later.*
- B. Hij heet toch Pepijn de Lom?  
*But isn't he called Pepijn de Lom?*

### 6.2.2 Participants

Recordings were made in five locations along the Dutch coast, covering four dialect groups (see Figure 1). We recorded Zeelandic Dutch in Zuid-Beveland (ZB), South Hollandic and North Hollandic in Rotterdam (RO) and Amsterdam (AM), West Frisian in Grou (GR) and Low Saxon in Winschoten (WI). We also recorded Standard Dutch speakers (SD).



Figure 1. Recording locations in the Netherlands.

As Table 1 shows, we recorded between 18 and 23 speakers of each variety, aged between 14 and 49. Participants were university students (SD), secondary school students (ZB), members of a Scouting club (RO, AM) or members of the local community (GR, WI). The speakers from Zuid-Beveland, Grou, and Winschoten were bilingual with Standard Dutch and their local language. All regional speakers and at least one of their parents were raised in the selected place and spoke the indigenous variety fluently. For Standard Dutch, the procedure was different, as the area where this variety is spoken is less determined by geographical boundaries. Speakers could participate if they reported to speak Standard Dutch. Besides self-reporting, two Dutch phoneticians independently judged each recording. Recordings were included if the judges agreed that the geographical and linguistic origin of the participants could not be determined by their accent. Except for the speakers of West Frisian and Standard Dutch, our speakers were less familiar with their local language as a written language, which may have had a negative influence on the fluency of the speech in the reading task of some speakers.

Table 1. Number, gender and average age (range) of speakers from Standard Dutch (SD), Zuid-Beveland (ZB), Rotterdam (RO), Amsterdam (AM), Grou (GR) and Winschoten (WI).

	female	Male	total	average age	age range	average age female	average age male
SD	14	9	23	23	18-30	23	23
ZB	8	10	18	19	16-33	21	17
RO	7	12	19	22	16-31	20	23
AM	7	11	18	31	14-45	33	30
GR	20	3	23	28	17-43	28	35
WI	14	4	18	32	15-49	30	41

Participants' recordings were excluded if they were (highly) disfluent or appeared to the experimenter not to speak naturally, if the speakers afterwards reported that they were dyslectic or had hearing problems, or if the speakers turned out not to satisfy the requirements with respect to their linguistic and/or geographical background. All participants were naive as to the purpose of the task and were paid for their participation.

### **6.2.3 Recordings**

The mini-dialogues were presented in a booklet, one dialogue per page. To prevent order effects, the dialogues were presented in pseudo-randomized order, which was reversed for half of the subjects per variety. We added 49 mini-dialogues from other experiments as fillers. Our speakers were recorded in pairs, to limit the effects of the experimenter's presence on dialect level. One speaker read the context sentence and the other the carrier sentence. The participants switched roles at the end of the task after they had repeated any mispronounced sentences. The Standard Dutch recordings were made in a professional studio at Radboud University Nijmegen; recordings of the local varieties were made in a quiet room either in the homes of our speakers or in a public building. We used a portable digital recorder (Zoom H4) with a 48 kHz sampling rate, 16 bit resolution and stereo format. The participants wore head-mounted Shure WH30XLR or Sennheiser MKE2 wired condenser microphones.

## **6.3 Nuclear tone preferences in different pragmatic contexts**

### **6.3.1 Labels**

One realization of each sentence by each speaker was included in the corpus we used for establishing NTFPs. If speakers produced the same sentence more than once with different nuclear contours, we chose the contour which was most frequently used by other speakers of the same variety. Utterances with irregular pitch patterns (e.g., creak) or with the nuclear pitch accent on the wrong word were excluded. The first author carried out the labelling, on the basis of an auditory impression of the speech waveforms and a visual inspection of the pitch curve and the narrowband spectrogram. Target sentences were initially annotated with the help of the pitch accent and boundary tone labels in Table 2 (Gussenhoven et al. 1998-2003, Gussenhoven 2005). The labelling proved to be unproblematic, with the exception of the IP-final high rise ( $H^* H\%$ ), the low rise ( $L^* H H\%$ ) and the low low rise ( $L^* H\%$ ), reported by Haan (2002:111-112) and Gussenhoven (2005). The low rise is characterized by a low-pitched accented syllable, contrasting with a rising or mid-pitched accented syllable in the high-rise, both being followed by mid pitch until a further rise on the last syllable. In a semantic task in which listeners had to rate degree of surprise, Gussenhoven and

Rietveld (2000) showed that the low rise and the high rise must be interpreted as belonging to different phonological contours. The low low rise has a steep final rise after a low level stretch from the accented syllable. However, on our IP-final syllables, phonetic rises were often hard to categorize in terms of the three rise types. The least problematic were contours that began early and rose late, which were assigned to L\* H%. Due to the short syllable rime on which the rises were produced, it often proved difficult to decide between L\*H H% and H\* H% in the remaining cases. For this reason, in the diagrams below the rises have been pooled.

Table 2. *Standard Dutch nuclear contours*<sup>2</sup> (Gussenhoven 2005).

<i>pitch accent</i>	<i>boundary tone</i>	<i>Description</i>
H*L	L%	fall
H*L	H%	fall-rise
H*L	0%	half-completed fall
H*	H%	high rise
H*	0%	level high
L*H	H%	low rise
L*	H%	low low rise
!H*L	L%	downstepped fall
!H*L	H%	downstepped fall-rise
L*HL	L%	delayed fall
L*HL	0%	delayed half-completed fall

## 6.3.2 Results

### 6.3.2.1 Statements

In all six varieties, statements were almost exclusively pronounced with a falling melody on the nuclear accented word, regardless whether that word occurred in IP-final or IP-medial position. However, speakers used a number of different falling contours. Besides the neutral contour (H\*L L%), we observed instances of falls with downstep (!H\*L L%), late falls (L\*HL L%) and half-completed falls with and without a late peak (H\*L 0% / L\*HL 0%). Interestingly, the distribution of fall types not only varied between final and non-final statements within varieties, but also between the varieties.

As the pie charts in Figure 2 illustrate, the largest proportion of non-standard falls could be observed in Amsterdam. They also illustrate that generally, IP-final falls showed more variation than IP-medial falls. The proportion of neutral falls is higher in IP-medial position, whereas Downstep is used more often for IP-

<sup>2</sup> Gussenhoven (2005) lists twelve other well-formed nuclear pitch contours of Dutch. They are not repeated here since they are not referred to in this paper.

final falls. Although the carrier sentences were short, in a number of cases speakers produced prenuclear accents. Prenuclear accents on words preceding IP-medial pitch accents were rare, which is why !H\*L L% is virtually confined to IP-final target words. Speakers from the western conurbation, as represented by the varieties of Rotterdam, Amsterdam and Standard Dutch<sup>3</sup>, applied Downstep more often (15% of IP-final falls) than the peripheral varieties of Zuid-Beveland in the South-East, or Grou and particularly Winschoten in the North (7% of IP-final falls). There was a significant association between the location of the speakers (Western conurbation or peripheral localities) and the frequency of Downstep  $\chi^2(1) = 7.31, p < .01$ . Odds ratios for speakers using Downstep were 2.26 times higher for Western speakers than for speakers in the periphery.

A second difference between final and non-final statements concerns the use of late falls and half-completed falls. Figure 2 shows that these were absent in most varieties on IP-final target words, but occurred in all varieties on IP-medial target words to various degrees. This IP-internal use of late or half-completed falls is rare in Standard Dutch, Zuid-Beveland, Grou and Winschoten, but did occur more frequently in Rotterdam (10%) and Amsterdam (47%). We postpone discussion of the phonetics of these falls to section 6.4.

### 6.3.2.2 Yes/no-questions

The pie charts in Figure 3 show for each variety which nuclear contours were used to express yes/no-questions and what their relative frequency was. Strikingly, Zuid-Beveland stands out from the others in having 81% falling nuclear contours in yes/no-questions. Of the other dialects, only nearby Rotterdam uses falls in yes/no questions (10%). The H%-ending contours outside Zuid-Beveland are divided over a rising nuclear melody and a falling-rising one. The latter melody is used significantly more often by speakers of Standard Dutch, Amsterdam, Grou and Winschoten than by speakers of Zuid-Beveland and Rotterdam (26% vs. 7%),  $\chi^2(1) = 22.19, p < .001$ . Based on the odds ratio, speakers of Standard Dutch, Amsterdam, Grou and Winschoten were 4.43 times more likely to use H\*L H% for IP-final yes-no-questions than speakers from Zuid-Beveland and Rotterdam. The category ‘other’, which includes those phonological contours that were produced in less than 5% of all utterances, is largest for Standard Dutch and Amsterdam, where it mainly consists of different fall types (neutral, downstepped and late falls). The phonetics of these falls will be discussed in section 6.4.

---

<sup>3</sup> In this paper, we treat Standard Dutch as part of the western conurbation because of its historically close relations to western varieties such as those spoken in Amsterdam and Rotterdam. Note, however, that our SD speakers originate from a variety of locations in the Netherlands.

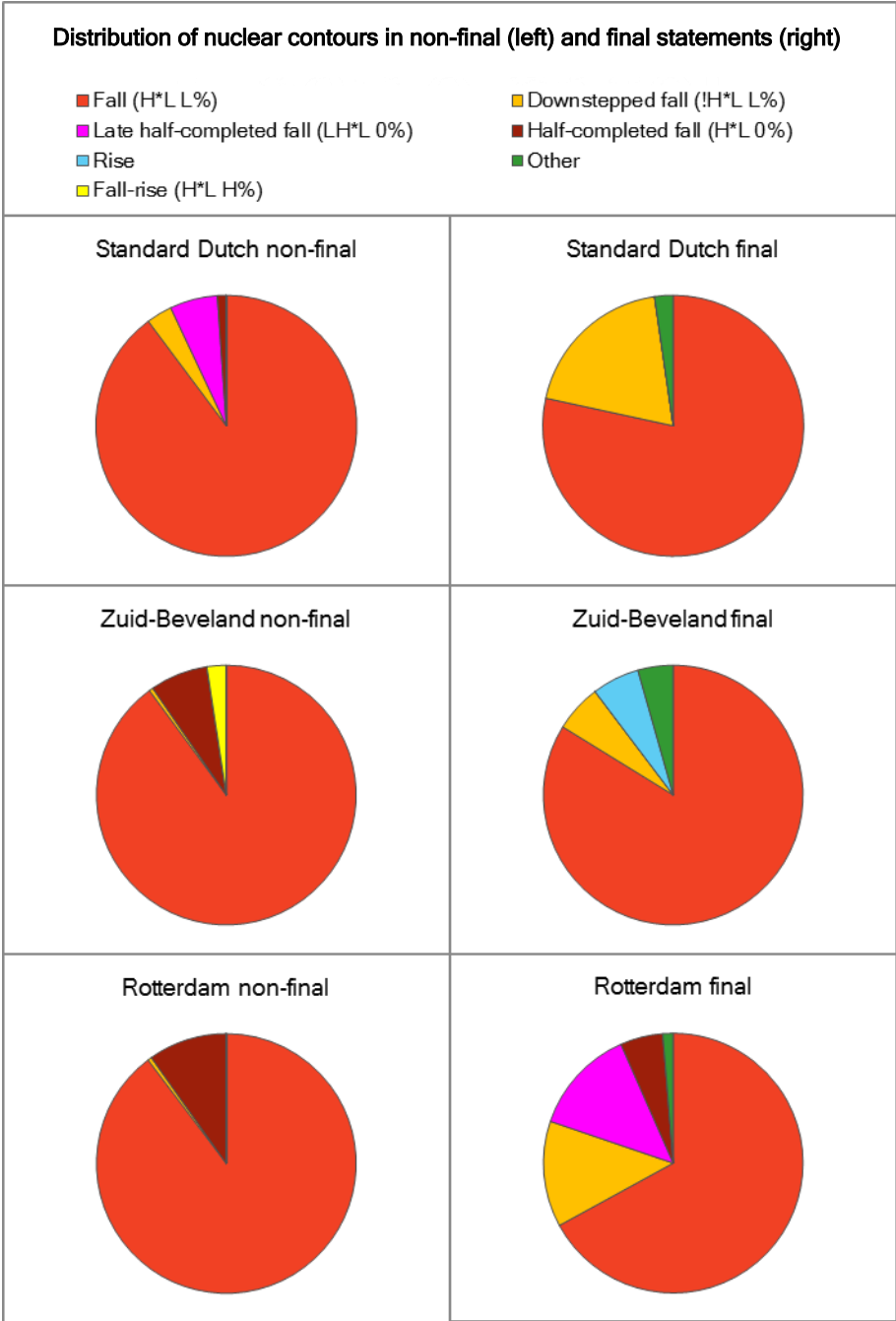


Figure 2. Distribution and relative frequency of nuclear contours in non-final statements, excluding mispronunciations. The category other consists of all nuclear contours that were produced in less than 5% of all utterances. [Continued on next page.]



[Figure 2. Continued.]



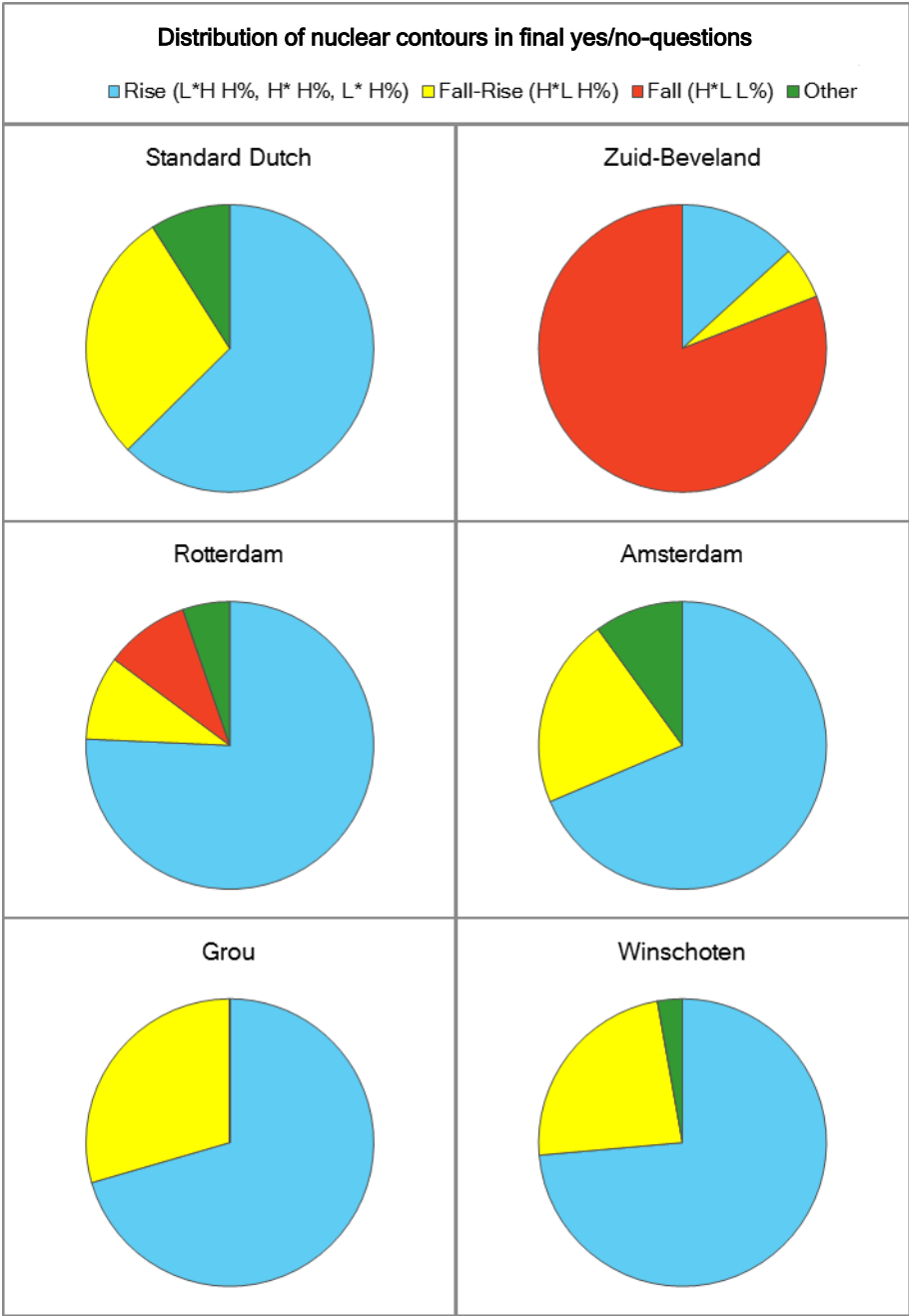


Figure 3. Distribution of nuclear pitch contours in yes/no-questions, excluding mispronunciations. The nuclear contours that are categorized as *Rise* in this figure include *low rises*, *high rises*, and *low low rises*. The category *Other* consists of all nuclear contours that were produced in less than 5% of all utterances.

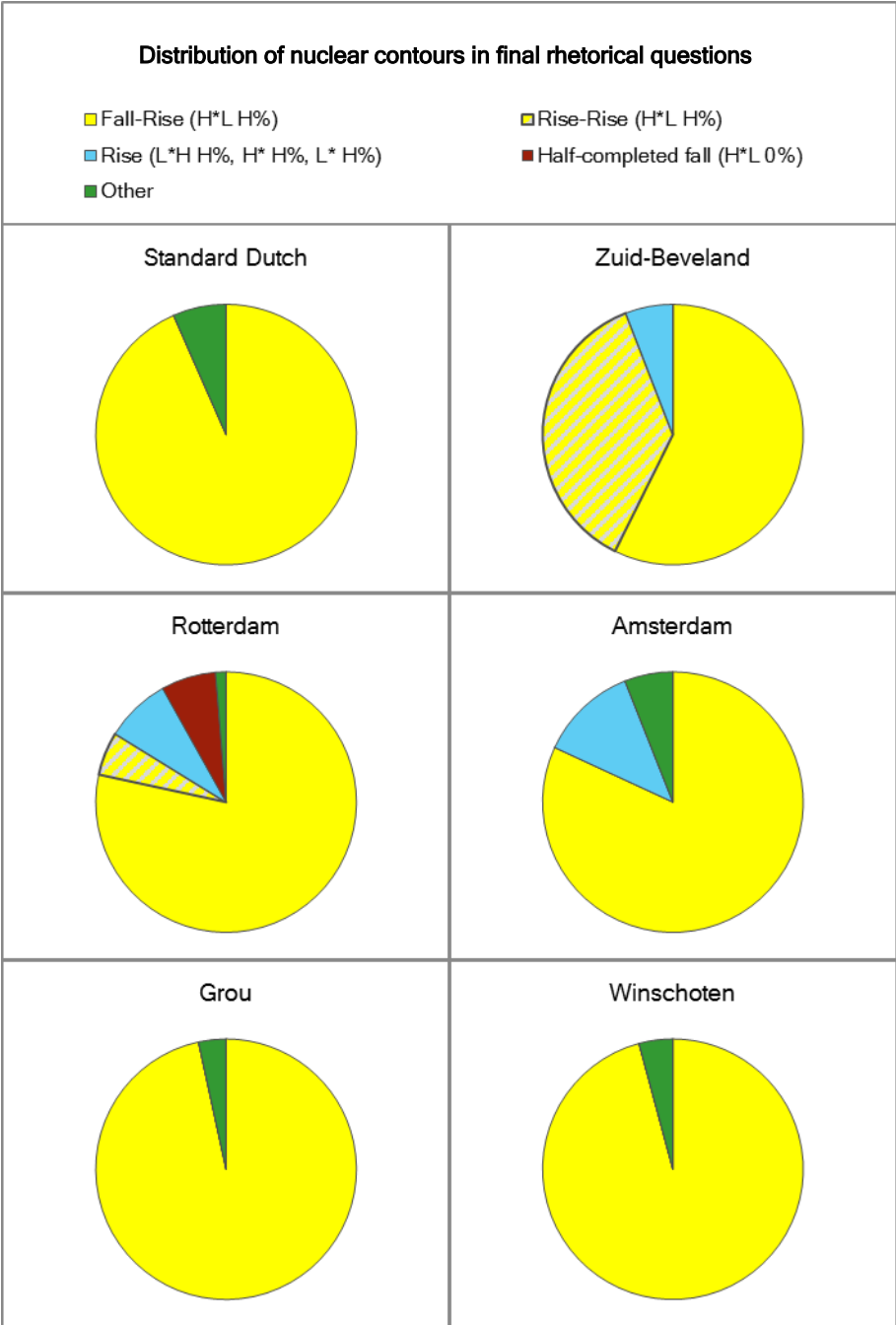


Figure 4. Distribution of nuclear pitch contours in rhetorical questions, excluding mispronunciations. The category *Other* consists of all nuclear contours that were produced in less than 5% of all utterances.

### 6.3.2.3 Rhetorical questions

As the pie charts in Figure 4 illustrate, speakers of all varieties used the fall-rise ( $H^*L\ H\%$ ) for the vast majority of rhetorical questions, varying from 84% in Rotterdam to 97% in Grou. A nuclear rise was used in over 5% in Zuid-Beveland, Rotterdam and Amsterdam. Speakers of Rotterdam used the half-completed fall ( $H^*L\ 0\%$ ) in 7% of the cases. Nuclear contours that were used in less than 5% of the rhetorical questions (the category *Other* in the pie charts) included neutral falls, half-completed falls and half-completed rises.

Although the six varieties showed a clear preference for fall-rises, the contour was not always realized in the same way. The depth of the valley between the nuclear peak ( $H^*$ ) and the boundary peak ( $H\%$ ) varied considerably between the six varieties. Averaged across male and female speakers, the excursion of the falling movement ranged from 1.3 semitones in Zuid-Beveland to 5.9 semitones in Winschoten. Zuid-Beveland speakers produced this shallow fall-rise in only 57% of all utterances. In 37%, they realized a phonetic variant of  $H^*L\ H\%$  that had little ( $< 5\text{ Hz}$ ) or no  $f_0$  drop between  $H^*$  and  $L$ , causing the contours to look like a sequence of rises. Although we occasionally also observed this ‘*rise-rise*’ in Rotterdam, this variant is clearly typical of Zuid-Beveland in our data. Section 6.4.2 gives a more detailed account of this contour.

### 6.3.3 Discussion

All nuclear contours observed in the six varieties are known to exist in Standard Dutch (see section 6.3.1). Table 3 lists the nuclear contours which were observed in each variety, including the three rises. It is stressed that empty cells indicate that a contour was not attested in our data, which does not imply that the contour does not occur at all in the variety concerned. In fact, all the ‘missing’ nuclear contours were found in all the dialects in Peters and Gussenhoven (ms), who recorded the same set of speakers, but whose reading materials included a wider variety of sentence types<sup>4</sup>. A case in point is the high rise ( $H^*\ H\%$ ) in Zuid-Beveland, which is used in 15% of all nuclear contours in Peters and Gussenhoven. This suggests that more rises in yes-no questions might have been used by our speakers in different contextualizations.

---

<sup>4</sup> Standard Dutch was not included in the analysis in Peters and Gussenhoven (ms). However, the four missing Standard Dutch nuclear contours are reported in Gussenhoven’s (2005) account of Standard Dutch intonation.

Table 3. *Overview of all possible and observed (‘✓’) nuclear contours per variety.*

<i>Nuclear contours</i>	<i>SD</i>	<i>ZB</i>	<i>RO</i>	<i>AM</i>	<i>GR</i>	<i>WT</i>
H*L L%	✓	✓	✓	✓	✓	✓
H*L H%	✓	✓	✓	✓	✓	✓
H*L 0%	✓	✓	✓	✓	✓	✓
H* H%	✓		✓	✓	✓	
H* 0%			✓	✓		
L*H H%	✓	✓	✓	✓	✓	✓
L* H%	✓		✓	✓	✓	✓
L*HL L%	✓	✓	✓	✓	✓	
L*HL 0%			✓	✓		
!H*L L%	✓	✓	✓	✓	✓	✓
!H*L H%				✓		

It is reasonable to conclude on the basis of Table 3 and the charts in Figure 2, 3, and 4 that these varieties share a tonal grammar and employ the same tonal contrasts, but that the distribution of nuclear contours varies systematically over pragmatic conditions, as reflected by their different frequencies within each of our three pragmatic conditions. Below we sum up the main findings and compare them to the results of Peters and Gussenhoven (ms) and to earlier findings in the literature on Standard Dutch.

First, statements were produced mainly with H\*L L% in all varieties, which is in line with findings for Standard Dutch (’t Hart, Collier, and Cohen 1990, Gussenhoven 2005, Haan 2002:115). Its use was most frequent in the Northeast (Grou and Winschoten). H\*L L% was additionally used for yes/no-questions by speakers of Zuid-Beveland, and to a lesser extent in Rotterdam. Peters and Gussenhoven (ms) similarly report a more frequent use of falling contours in the northeastern varieties than in Zuid-Beveland and the central varieties. The use of H\*L L% for questions, which in their study included several question types such as polar questions and wh-questions, was reported for all varieties, but was considerably more frequent in Zuid-Beveland. Apparently, whereas H\*L L% has both a declarative and a strong interrogative meaning in Zuid-Beveland, its interrogative value is weaker in the northeastern varieties, in which rising intonation is preferred for questions.

The association between questions and H%-ending contours is common (cf., e.g., Bolinger 1978). For Standard Dutch questions, Haan (2002:112) showed that H\*L H% is the most common tune for yes/no-questions with the nuclear accented word in IP-medial position (72%). In our IP-final data, simple rises are by far the most frequently used intonation contour for yes/no-questions (70%) in all varieties except Zuid-Beveland. Following Lickley, Schepman and Ladd (2005), who found that out of the nuclear contours ending with a rise (H\*L H% and simple rises) the simple rises are used more frequently in IP-final position (95%) than in IP-non-final position (44%) in conversational speech by seven

speakers recruited from the Nijmegen student population, we explain the avoidance of the more complex contour on sentence-final monosyllables as a response to the time pressure (cf. Hanssen, Peters, and Gussenhoven ms).

Our results reveal a substantial use of late falls by speakers of Rotterdam and particularly Amsterdam. Peters and Gussenhoven (ms) similarly observe that the late fall is used most frequently in Amsterdam, followed by both Grou and Rotterdam. They also report that late falls were always found on the same part of the narrative, suggesting that it reflects a similar meaning across varieties. The late peak contours are further discussed in section 6.4.1.

As for rhetorical questions, the dominant use of H\*L H% in all varieties is in line with earlier findings for Standard Dutch. The persuasive use of the Dutch modal particle *toch* ‘after all, surely’ will typically trigger a falling-rising melody on the corrective word. Van de Ven and Gussenhoven (2011) found for nine Standard Dutch speakers that they consistently produced the H\*L H% contour on non-IP-final nuclear syllables in contextualized sentences with *toch*.

To conclude, our results suggest, first, that the five regional varieties share their tonal grammar and tonal contrasts with Standard Dutch. Second, they show that the varieties vary semantically, as evidenced by the variation in the distribution of nuclear contours over sentence types and in the frequency with which they are used. In section 6.4, it is shown that they also vary phonetically in non-trivial ways.

## 6.4 Dialect-specific contour realizations

In this section, three dialect-specific nuclear contours are described, none of which occur in the standard language:

1. falling nuclear contours in Rotterdam Dutch and Amsterdam Dutch, which are used in statements (6.4.1);
2. declarative and interrogative falling nuclear contours in Zuid-Beveland (6.4.2);
3. the Zuid-Beveland ‘rise-rise’ (6.4.3).

For each of these dialect-specific intonation contours, we will explore whether we are dealing with phonological or phonetic (i.e., realizational) variation.

### 6.4.1 Falls in Rotterdam and Amsterdam

As noted in section 6.3, Rotterdam and Amsterdam speakers showed more variation than the other speakers in their choice of contour for statements by using a type of fall which appeared as a truncated fall in IP-final position and as

a slowly descending fall in IP-medial position. There are a number of initially plausible phonological interpretations of the IP-medial slow or late falls in these dialects. First, Standard Dutch has been described as having a delayed fall, which is analyzed as the standard H\*L L% contour with a prefixed L\*, giving L\*HL L%. Another is the half-completed fall, H\*L 0%, while a third is a delayed half-completed fall, i.e. the half-completed fall with prefixed L\*. Examples of these contours from the ToDI interactive course (Gussenhoven et al. 2006-2010) are given in Figure 5. The delayed H\*L L% literally has a delayed peak, but is otherwise quite like the standard sharp fall (panel a). It is auditorily rather different from the shallow falls of Amsterdam and Rotterdam. The H\*L 0% in panel (b) is quite like the shallow falls of Amsterdam and Rotterdam and is auditorily very similar. Late peaks for this contour are described with prefix L\*, i.e., L\*HL 0% (panel c). In neither location did we find falling contours that combined late peaks with steep falls. Late peaks were either truncated (IP-finally position) or followed by a slowly descending pitch (IP-medially). This means that the variation in our data is between plain half-completed falls (H\*L 0%) and delayed half-completed falls (L\*HL 0%).

The half-completed fall in our data has a shallow falling slope after the nuclear peak, which does not reach the bottom of the speaker's range. Phonologically, it lacks a L% boundary tone. The delayed half-completed fall is characterized in these data by a peak occurring late in the IP-final syllable or, in IP-medial position, just inside or after the accented syllable where it is followed by low or slowly descending pitch. As the pie charts in Figure 2 show, the contours without final L% are considerably more frequent in the Amsterdam data than in the Rotterdam data (44% vs 12%).

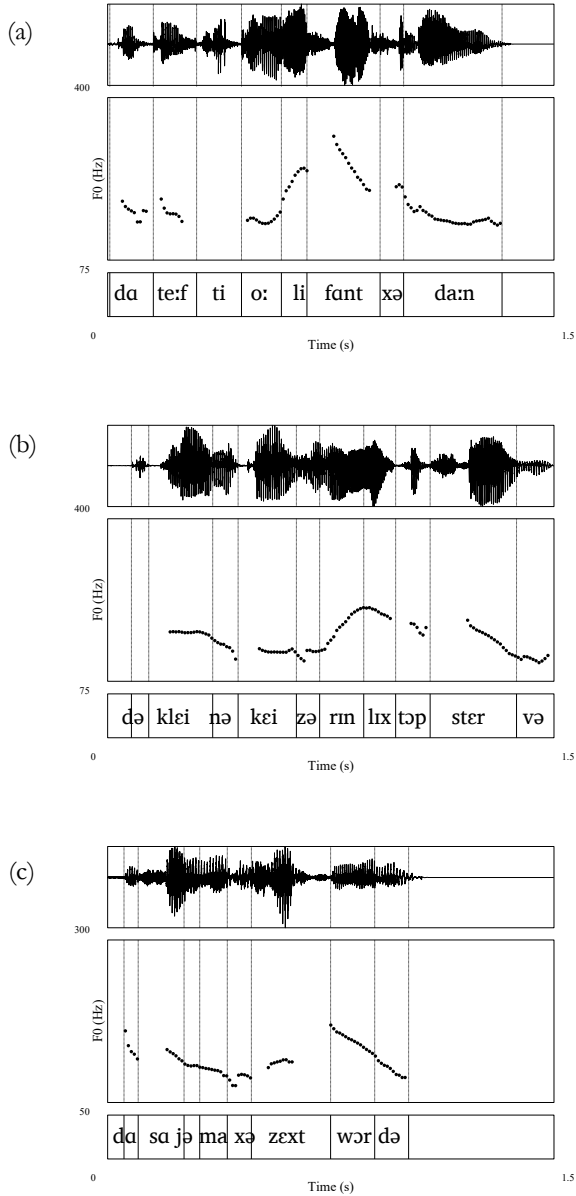


Figure 5. F0 tracks of (a) L\*HL L% on *Dat heeft die olifant gedaan* 'The elephant did that' (ToDI Exercise 6, nr 5), female speaker, (b) H\*L 0% *De kleine keizerin ligt op sterven* 'The little empress is dying' (ToDI section 1.4), male speaker, and (c) L\*HL 0% *Dat zal je maar gezegd worden!* 'Imagine being told that!', approximately 'The cheek!' (ToDI section 6), male speaker.

IP-final falls

Panels (a) and (b) of Figure 6 show the (late) half-completed fall and the neutral fall in the data for Rotterdam and Amsterdam. They are averaged  $f_0$  measurements contours labeled  $L^*HL\ 0\%$  and  $H^*L\ 0\%$  on the one hand (blue line), and all  $f_0$  contours labeled  $H^*L\ L\%$  on the other (dotted green line), over all speakers. Examples from individual Amsterdam speakers are given in Figure 7. A neutral fall is presented in panel (a), whereas panels (b) and (c) give two examples of IP-final late falls. Because of truncation in IP-final position, the late fall may in fact visually resemble a rising movement rather than a fall, but is auditorily very different from a rise (panel c).

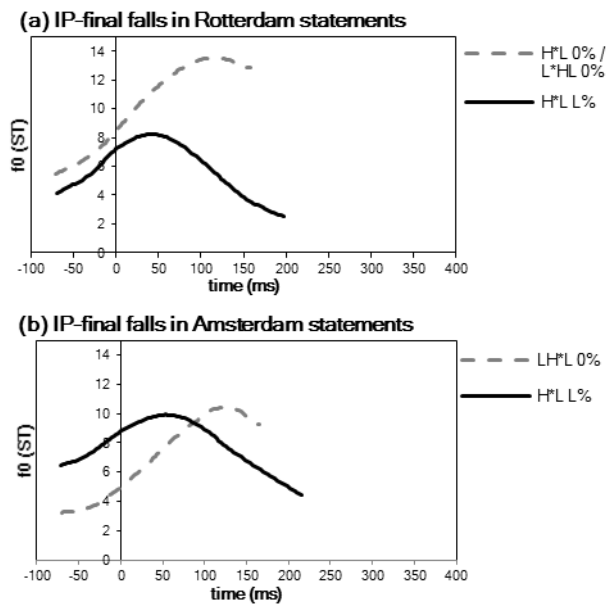


Figure 6. Averaged  $f_0$  contours (in semitones re 100 Hz) of IP-final fall types. Rotterdam (panel a) and Amsterdam (panel b). The averaged contours represent the nuclear accented monosyllabic word (e.g., lom) (Rotterdam, N=51 ( $H^*L\ L\%$ ) and N=14 ( $H^*L\ 0\% / L^*HL\ 0\%$ ); Amsterdam, N=33 ( $H^*L\ L\%$ ) and N=23 ( $L^*HL\ 0\%$ )). The Y-axis is placed at the beginning of the nuclear vowel.



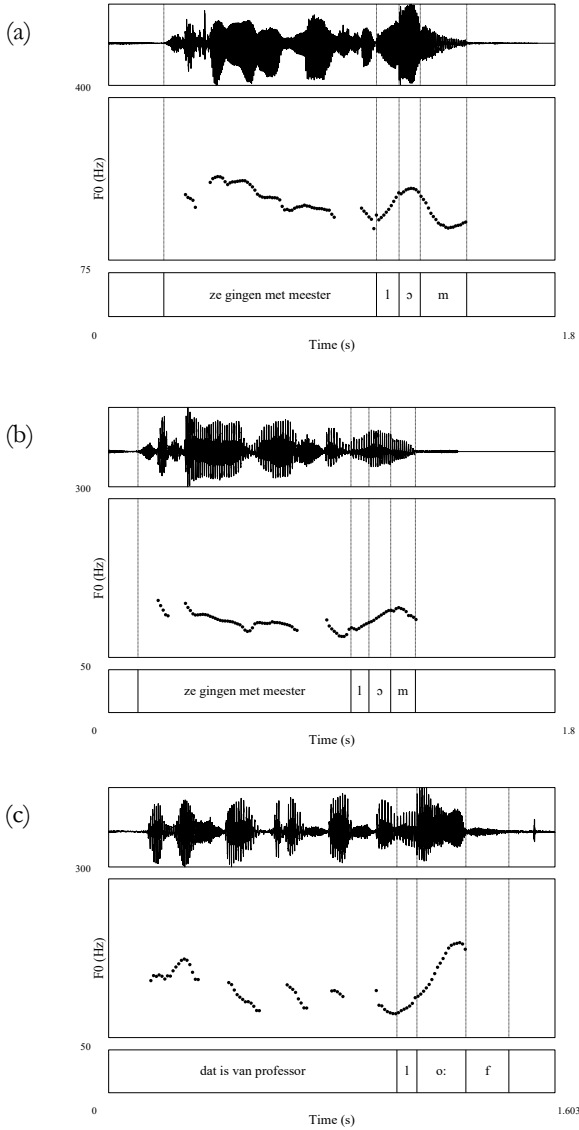


Figure 7. F0 tracks of *Ze gingen met meester Lom* ‘They went with mister Lom’ produced in Amsterdam with (a) H\*L L%, female speaker and (b) L\*HL 0%, male speaker. Panel (c) shows an F0 track of L\*HL 0% on *Dat is van professor Loof* ‘That belongs to Professor Loof’, male speaker.

In both Rotterdam and Amsterdam, the peak in late falls occurs about 70 ms later in the syllable than in neutral falls. Measured as a proportion of the sonorant rime duration, which is the duration of the vowel in *lof/loof* (/ɔ/ and /o:/), and the duration of the vowel plus coda consonant in *lom/loom* (/ɔm/ and /o:m/), the

late peak occurs about 70-75% into the sonorant rime, versus 20-25% for the peak in neutral falls. While there is only little overlap, the individual speakers' peak timings do not cluster tightly around the 25 and 75% marks, but cover the entire first and second half of the scale, as is illustrated in Figure 8 for both Amsterdam and Rotterdam. Besides peak timing, the two contours differ in f<sub>0</sub> excursion after the peak due to truncation of the later peaks. Unexpectedly, falls with late peaks occur more often in target words with shorter sonorant rimes (i.e., *lof/loof* as opposed to *lom/loom*). This is also reflected in the averaged contours in Figure 6, where the blue line representing the (late) half-completed falls is around 30-40 ms shorter (i.e., shorter rime duration) than the line representing the neutral fall. In other words, the shorter line is based on proportionally more short-rime pitch accents.

Figure 6 shows a considerable difference in peak scaling between the earlier and later peaks in the Rotterdam data, which is to be attributed to different proportions of speaker gender. The difference in scaling between male and female peaks is  $\pm 9.5$  ST for neutral falls and  $\pm 7.7$  ST for (late) half-completed falls. Since 43% (N=6) of (late) half-completed falls, but only 27% (N=14) of neutral falls were produced by a female speaker, f<sub>0</sub> is relatively low for the averaged H\*L L% contour in Figure 6.

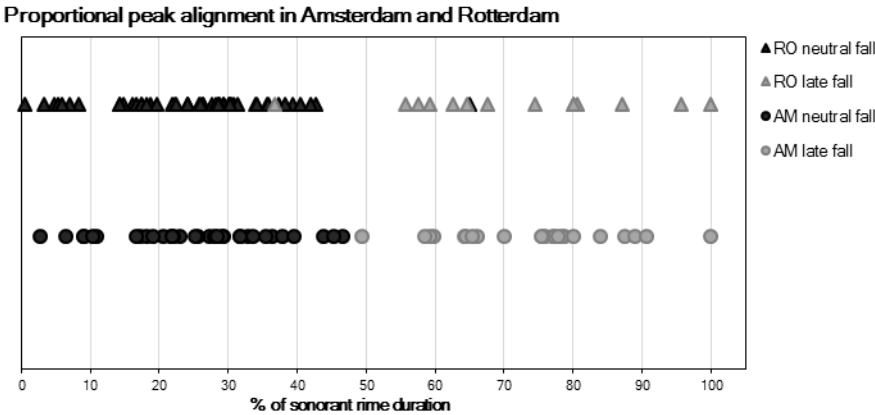


Figure 8. Scatterplot of proportional peak alignments in Amsterdam (circles) and Rotterdam (triangles). IP-final neutral falls are black; late (half-completed) falls are gray.

### IP-medial falls

In both the Amsterdam and Rotterdam varieties, falls in statements with IP-medial target words were realized with a falling nuclear contour in almost all utterances. There was greater variation in these contours in the Amsterdam

group than in the Rotterdam group. Some falls can be characterized as standard  $H^*L\ L\%$ , with their steep falls and low level or slightly declining post-nuclear stretches. Other falls descended more slowly, either to low pitch, as in the case of neutral falls, or to mid pitch. In addition to the shape of the downward slope, the contours varied in the location of the peak, which could be inside the nuclear accented syllable or after it, sometimes in the initial syllable of the next word. The averaged graphs in Figure 9 illustrate the shape of IP-medial, non-final (late) half-completed falls together with that of the neutral falls in Rotterdam (panel a) and Amsterdam (panel b). Each contour represents the nuclear fall from the beginning of the accented word to the end of the first postnuclear word (e.g., *MANderen willen*). A comparison between the Amsterdam and Rotterdam contours above all shows their similarity, the main difference lying in the timing of the peak of the (late) half-completed fall.

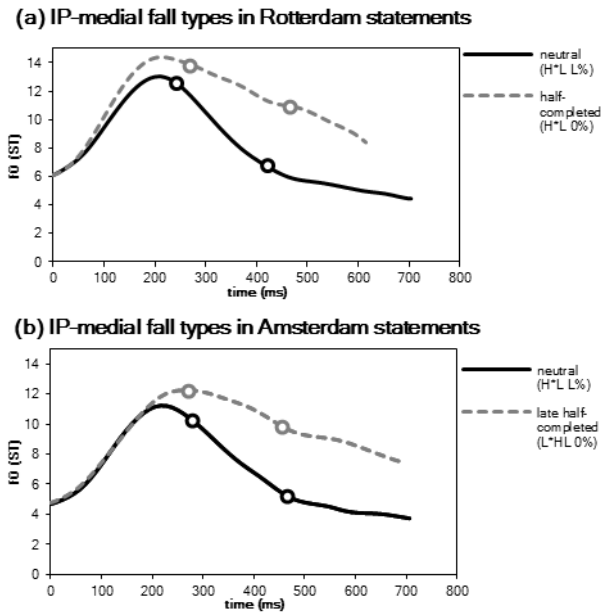


Figure 9. Averaged  $f_0$  contours in semitones of IP-medial fall types in Rotterdam (panel a) and Amsterdam (panel b). The averaged contours represent the nuclear accented word and the first postnuclear word (Rotterdam:  $H^*L\ L\%$   $N=202$ ,  $H^*L\ 0\%$  /  $L^*HL\ 0\%$   $N=22$ ; Amsterdam:  $H^*L\ L\%$   $N=107$ ,  $L^*HL\ 0\%$   $N=96$ ). Open circles indicate the endings of the accented syllable and the beginnings of postnuclear words.

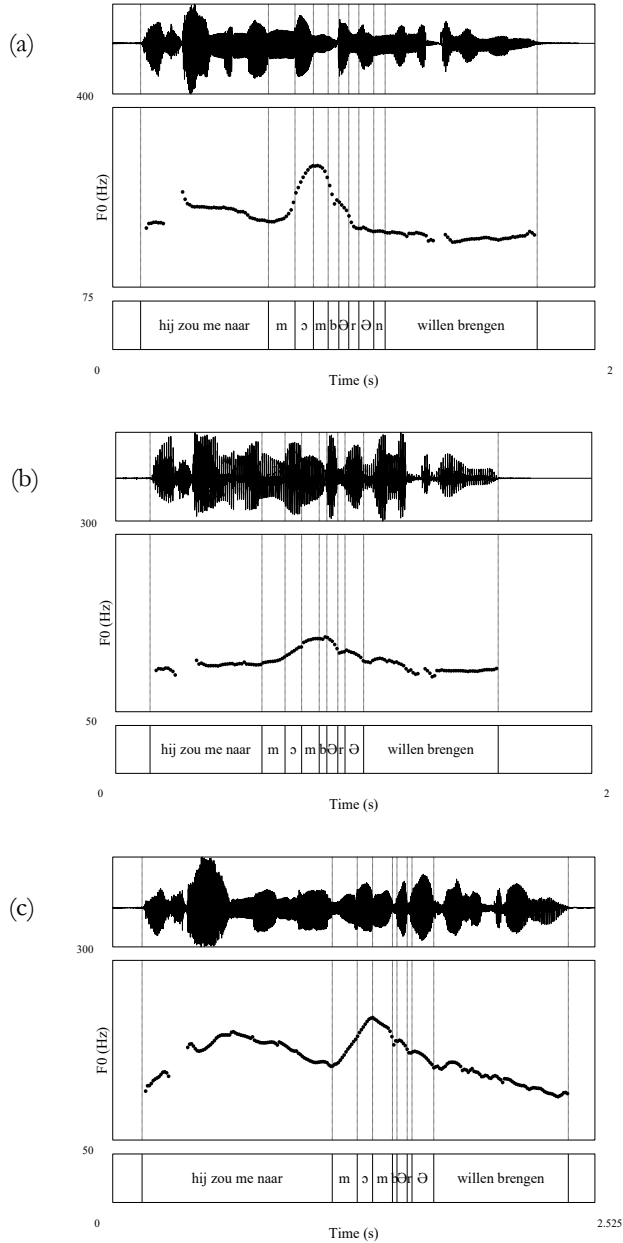


Figure 10. F0 tracks of *Hij zou me naar MOMberen willen brengen* 'He wanted to take me to Momberen' produced in Amsterdam with (a) H\*L L%, female speaker, (b) L\*HL 0%, male speaker, and (c) H\*L 0%, male speaker.

Figure 10 presents three fall types observed in Amsterdam. Panel (a) shows an early peak contour with a rapid fall inside the accented word; panel (b) shows a late-peak contour with slowly descending pitch after the peak. In panel (c), a normal peak is combined with slowly falling pitch. The question arises whether the late peak contour differs phonologically from the neutral fall. One argument for considering them as separate phonological contours is that rapid falls do not combine with late peaks in either location. On the other hand, the wide range that is covered by the data points representing proportional peak timing in Figure 8 pleads for a phonetic interpretation of the data. In this view, the fall shapes that we have observed in Rotterdam and Amsterdam are phonetic variants of the same phonological category. The varieties' difference with the standard language lies in the greater degree of variation in the Amsterdam and Rotterdam fall shapes.

#### 6.4.2 Zuid-Beveland declarative and interrogative falls

Zuid-Beveland speakers were exceptional in their use of the falling contour (H\*L L%) in yes/no-questions. Figure 11 compares averaged intonation contours of IP-final nuclear falls as used in statements (the declarative fall) and yes/no-questions (the interrogative fall) by the same set of speakers (N=11). The interrogative fall was pronounced with higher overall pitch, had a later peak, and a larger and steeper falling excursion than the declarative fall.

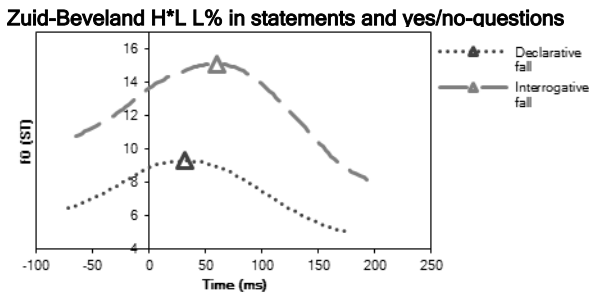


Figure 11. Averaged  $f_0$  contours in semitones of Zuid-Beveland declarative falls and interrogative falls. The beginning of the nuclear vowel is set to 0 ms. The triangles represent the location of the nuclear pitch peak.

For each speaker and each sentence type, we measured a number of acoustic variables in and around the IP-final test word that carries the nuclear pitch accent. We measured the mean duration of the sonorant rime (SONRIME), the mean duration, excursion and slope of the falling movement from H\* to L (FALLDURATION, FALLEXCURSION and FALLSLOPE), the mean timing of the peak, both relative to the beginning of the nuclear vowel (PEAKTIMING) and as

a proportion of the sonorant rime duration (PROPPEAKTIMING), and the mean scaling of the peak (H-SCALING) and of the following elbow (L2SCALING), as averaged over the four words *lof*, *loof*, *lom* and *loom*. We found that interrogative falls were realized with *longer* sonorant rime durations (mean  $\Delta$  12 ms), *larger* falling excursions (mean  $\Delta$  1.9 ST), *steeper* slopes (mean  $\Delta$  15 ST/s), *later* absolute and proportional peaks (mean  $\Delta$  23 ms and 12%, respectively), and *higher* peaks (mean  $\Delta$  4.5 ST) and *higher* valleys (mean  $\Delta$  2.6 ST), compared to declarative falls. The duration of the falling movement did not differ between interrogative and declarative falls. All data were submitted to several paired sample *t*-tests to establish which phonetic properties played a significant role in the distinction between declarative and interrogative falls. The results showed that Zuid-Beveland speakers realized the interrogative fall significantly differently from declarative falls in terms of SONRIME, FALLEXCURSION, FALLSLOPE, PEAKTIMING, PROPPEAKTIMING, H-SCALING and L2SCALING and that all variables except (proportional) peak timing showed high correlations. The results of the paired sample *t*-tests, along with mean values and standard deviations for declarative and interrogative falls, are presented in Table 4 for all dependent variables.

Table 4. Results of dependent *t*-tests comparing declarative falls (DF) and interrogative falls (IF).

		DF	IF	<i>t</i> -value	Proba- bility (p)	Correlation ( <i>r</i> ) plus probability
SON RIME	mean	179 ms	191 ms	$t(10) = 2.58$	< .05	.64*
	sd	19	16			
FALL DURATION	mean	116 ms	116 ms	$t(10) = -.10$	<i>n.s.</i>	.71*
	sd	19	14			
FALL EXCURSION	mean	5.2 ST	7.1 ST	$t(10) = 5.96$	< .001	.80**
	sd	1.3	1.7			
FALL SLOPE	mean	46.3 ST/s	61.3 ST/s	$t(10) = 4.35$	< .001	.61*
	sd	11.5	13.9			
PEAK TIMING	mean	35 ms	58 ms	$t(10) = 3.41$	< .01	-.06
	sd	15	15			
PROPPEAK TIMING	mean	19 %	31 %	$t(10) = 2.91$	< .05	-.43
	sd	8	8			
H- SCALING	mean	10.1 ST	14.6 ST	$t(10) = 9.52$	< .001	.97***
	sd	5.1	5.8			
L2 SCALING	mean	4.9 ST	7.5 ST	$t(10) = 5.90$	< .001	.97***
	sd	4.7	5.5			

To summarize the results of our comparison of Zuid-Beveland interrogative and declarative falls, we can say that the phonetic realization of nuclear falls is dependent on sentence type. In other words, Zuid-Beveland speakers systematically use phonetic variation to communicate pragmatic differences.

### 6.4.3 Zuid-Beveland ‘rise-rise’

Many of the nuclear melodies in rhetorical questions were pronounced as ‘rise-rises’ by the Zuid-Beveland speakers. These contours are here analyzed as phonetic variants of the fall-rise, or H\*L H% in which the valley has been smoothed out, a case of low target undershoot. The contours have an initial steep rise to H\* (similar to the rise preceding H\*L H%), which is followed by a shallower rise to the end, or by having an initial rise followed by a level part and then a final rise. Contours were labeled as ‘rise-rise’ variants of H\*L H% either if no f0 dip (L) could be labeled between the two high targets (H\* and H%) or if no final rise followed the f0 dip or level tone. Using this criterion, 37% of nuclear contours in rhetorical questions could be labeled as ‘rise-rises’ against 57% ‘fall-rises’. Figure 12 presents example contours for the words *lof* and *loom* in IP-final position. We can see that the ‘dipless’ variants (rise-rises in panels b and d) occur both on the shorter and longer target words. Moreover, the top right panel (b) shows that the final rise may be somewhat truncated. In case of extreme truncation, this may result in a contour that just levels off after H\*, illustrated in panel (e).

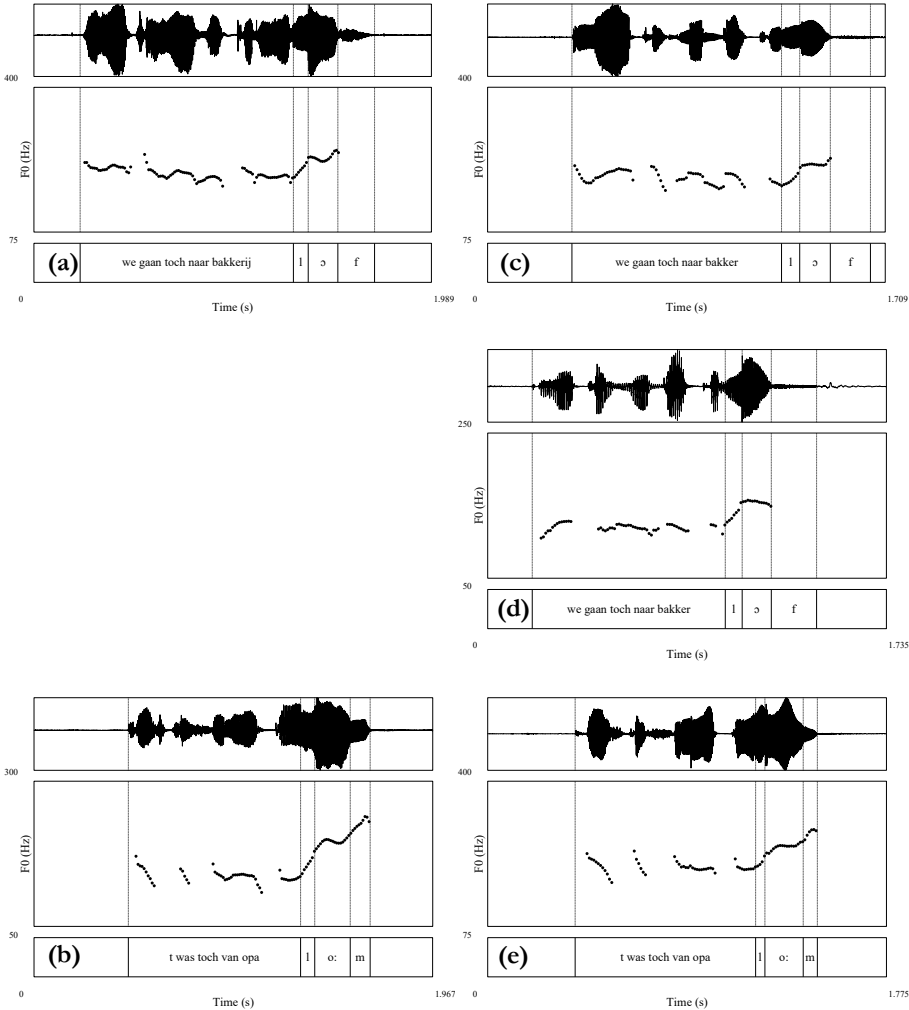


Figure 12. F0 tracks of Zuid-Beveland *Fall-Rise* (lefthand panels a and b) and *Rise-Rise* contours (righthand panels c, d and e) on the target words *lof* and *loom* in *We gaan toch naar bakker(ij)* *LOF?* ‘But aren’t we going to baker Lof?’ and *‘t Was toch van opa Loom?* ‘But didn’t it belong to grandfather Loom?’.

The argument for considering the rise-rise a phonetic variant of the fall-rise is that the variation between the clearest examples of the rise-rise and fall-rise is continuous. That is, there is no dichotomy between contours with and without a clear dip. To illustrate this, Figure 13 shows four graphs averaged over the four target words *lof*, *loof*, *lom* and *loom* which have been aggregated over the following sets: (a) all Zuid-Beveland H\*L H% contours (i.e., including both rise-rises and



fall-rises); (b) Zuid Beveland contours with an  $f_0$  drop of  $\geq 5$  ST ('fall-rises'); (c) Zuid Beveland contours with an  $f_0$  drop of less than 5 ST ('rise-rises'); (d) standard Dutch fall-rises. Duration differences reflect the fact that contour with  $f_0$  dips were produced more often on longer target words: the shorter the target word, the shallower the dip. The averaged contour of H\*L H% in Standard Dutch (d) is there to demonstrate that the Zuid-Beveland H\*L H% is comparatively flat even if 'rise-rises' are excluded from the comparison.

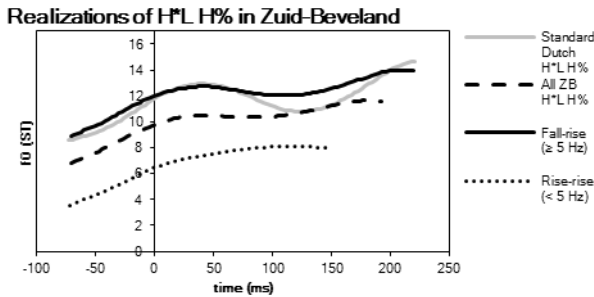


Figure 13. Averaged contours of all Zuid-Beveland H\*L H% contours, all Rise-Rise contours, all Fall-Rise contours, compared with all Standard Dutch H\*L H% contours. The ordinate is semitones re 100 Hz; the abscissa is time in ms. The beginning of the nuclear vowel is set to 0 ms.

A final observation to be made about the H\*L H% contour in Zuid-Beveland is that female speakers produced the 'fall-rise' (i.e., with an  $f_0$  dip and a final rise) significantly more often than male speakers: 86% for female speakers compared to 42% for male speakers, Pearson  $\chi^2(1) = 12.84$ ,  $p < .001$ . Based on the odds ratio, the chance that a 'fall-rise' is produced by a female speaker is 8.4 higher than that is produced by a male speaker.

#### 6.4.4 Discussion

We have presented three intonation contours which appeared to occur in specific varieties only and which may thus contribute to the dialectal signature of these varieties. First, in Amsterdam and Rotterdam shallow falls occur much more frequently than in the other varieties, the *half-completed fall* and the *delayed half-completed fall*, both of which were characterized by a slowly descending slope on IP-medial target words and by truncation on final target words. We did not observe a clustering of peak delays around two values and thus no clear distinction between the non-delayed and delayed versions of the half-completed fall. These (delayed) slowly falling rise-falls are evidently characteristic of the two dialects, in particular Amsterdam. This delayed fall is widely used in spoken advertisements for a general Dutch audience, and although such advertisements

normally don't relate in any way to Amsterdam or Rotterdam, its use may be motivated by the covert prestige of the Amsterdam accent (cf. Gussenhoven 1985). We did not find examples of the delayed neutral fall, i.e. delayed falls with a steeply descending slope, which the ToDI course (Gussenhoven et al. 1998-2003) says is used in the standard language "in a particular style of speech, which is sometimes used by speakers addressing children."

Next, we saw that Zuid-Beveland differed from other varieties in two ways. First, its speakers often used H\*L L% for yes/no-questions which differed phonetically from the same contour as used in declaratives. The interrogative fall is an enhanced version of the declarative fall. It is realized at a higher pitch level, with a larger pitch excursion, a steeper slope and a later peak. Differences in peak timing to signal variation between statements and questions have been reported for a number of languages. In Hungarian, Neapolitan Italian and Swedish, for example, questions are similarly associated with later and often higher peaks than statements (Gósy and Terken 1994, D'Imperio and House 1997, House 2003). For Hungarian and Neapolitan Italian, the alignment difference is claimed to be phonologized, whereas the pitch accents in Swedish statements and questions are interpreted as phonetic variants.

In Zuid-Beveland, interrogative falls are not only realized with later and higher peaks, but the overall pitch level is raised. The features can be explained in terms of the Frequency Code (Ohala 1983, also see Gussenhoven 2004: ch 4, Ladd 2008:126). Via a paralinguistic channel of communication, high pitch may signal 'uncertainty' and as such be associated with question intonation.

Besides the interrogative fall, speakers from Zuid-Beveland also differed from other speakers in their realization of the fall-rise, which was much shallower compared to the other varieties, often resulting in a 'rising-rising' intonation contour. This shape was interpreted as resulting from extreme undershoot of the L-target between the two peaks, which may be combined with truncation of the final rise on shorter words. The *rise-rise* and fall-rise are thus to be interpreted as phonetic variants of the same phonological contour, H\*L H%. Based on our data, it is not yet possible to say whether the flat realization of H\*L H% is typical of Zuid-Beveland in general or whether speakers only produce it in IP-final condition, as a response to time pressure. Therefore, it is desirable that more research is devoted to this issue, specifically by collecting data of IP-medial realizations.

From a sociolinguistic perspective, the fact that more female than male speakers in Zuid-Beveland produced the fall-rise with a dip is in line with findings that female speakers tend towards the standard more than males (Chambers and Trudgill 1998). We assume that dipless realization of the fall-rise is an example of language change 'from below', which must have spread from the urban center

of Rotterdam. On this view, the feature is strongly recessive in urban Rotterdam, but lingering on in the more peripheral Zuid-Beveland.

## 6.5 Conclusion

There are two broad conclusions to be drawn from our investigation into the distributional and phonetic differences between non-tonal varieties of Dutch. First, distributional and realizational differences can both contribute to a dialect's 'signature'. Based on our data set, which was restricted to IP-medial statements as well as IP-final statements, yes/no-questions and rhetorical questions, we have demonstrated a number of features determining these 'signatures'. These include the more frequent use of H\*L H% in questions in the northern three dialects as well as in Standard Dutch; and the frequent use of (late-peak) half-completed falls in Rotterdam and Amsterdam. It is of course possible that other sentence types, collected specifically for the purpose of finding dialect-specific semantic or systemic features of intonation, might bring more differences to the surface.

The second point is that the four ways in which intonation may vary according to Ladd (1996/2008) may not be adequate to describe dialectal differences in intonation. Ladd defines semantic differences as variation in the meaning of equatable contours. Dialectal differences in the frequency with which contours are used in the same pragmatic context are captured by this definition only if we allow intonational meaning to be defined by the distributional profile of a contour over some exhaustive set of pragmatic categories. If we were to proceed along that way, we would end up assuming that speakers of different dialects by definition express the same range of meanings in each of these pragmatic conditions. This uncomfortable circularity suggests that determining the communicative effects of intonation contours in different dialects by both native and non-native listeners should be a necessary follow up to research of the sort reported here.

Finally, it is stressed that our data are from secondary school students and similar subjects. Besides their young age, they tend to come from non-working class families. This has biased our research towards a more standard form of speech than is likely to exist in each of the six locations.

## References

- Bolinger, D. (1978). Intonation across languages. In: Greenberg, J. (Ed.), *Universals of human language* Vol. 2: Phonology. Stanford: Stanford University Press, 471- 524.
- Chambers, J.K. and Trudgill, P. (1998). *Dialectology*. Cambridge: Cambridge University Press.
- Daan, J. (1938). *Dialect and pitch pattern of the sentence*. Proceedings of the 3rd international congress of phonetic sciences, Ghent, 473-480.
- D’Imperio M. and House, D. (1997). Perception of questions and statements in Neapolitan Italian. In: Kokkinakis, G., Fakotakis, N., and Darmatas, E. (Eds.), *Proceedings of Eurospeech ‘97*, vol. 1. Rhodes. 1251-1254.
- Gooskens, C. (1997). *On the role of prosodic and verbal information in the perception of Dutch and English language varieties*. Wageningen: Ponsen en Looijen.
- Gosy, M. and Terken, J. (1994). Question marking in Hungarian: Timing and height of pitch peaks. *Journal of Phonetics* 22, 269-281.
- Grabe, E. (2004). Intonational variation in urban dialects of English spoken in the British Isles. In: Gilles, P. and Peters, J. (Eds.), *Regional variation in intonation*. Tübingen: Niemeyer, 9-31.
- Grabe, E. and Post, B. (2002). Intonational variation in the British Isles. In: Bel, B. and Marlien, I. (Eds.), *Proceedings of Speech Prosody 2002*, Aix-en-Provence, 343-346.
- Gussenhoven, C. (1985). ‘Intonation: A whole autosegmental language’. In H. van der Hulst and N. Smith (Eds.), *Advances in non-linear phonology*. Dordrecht: Foris. 117-131.
- Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge: Cambridge University Press.
- Gussenhoven, C. (2005). Transcription of Dutch intonation. In: Jun, S.-A. (Ed.), *Prosodic typology: The phonology of intonation and phrasing*. Oxford: Oxford University Press, 118-145.
- Gussenhoven, C. and Rietveld, A. (2000). The behavior of H\* and L\* under variations in pitch range in Dutch rising contours. *Language and Speech* 43, 183-203.
- Gussenhoven, C., Rietveld, T., Kerkhoff, J., and Terken, J. (1999-2003). *ToDI: Transcription of Dutch Intonation*. Second edition. <http://todi.science.ru.nl/>. Gussenhoven et al. (1998-2003).
- Gussenhoven et al. (2006-2010). ToDI interactive course.
- Haan, J. (2002). Speaking of questions. PhD thesis. Utrecht: LOT.
- Hanssen, J., Peters, J., and Gussenhoven, C. (ms). *Regional variation in phonetic responses to time pressure in Dutch IP-final nuclear contours*.
- House, D. (2003) Perceiving question intonation: the role of pre-focal pause and delayed focal peak. Proceedings of the 15th ICPhS, Barcelona, 755-758 .
- Ladd, D.R. (1996). *Intonational Phonology*. Cambridge: Cambridge University Press.
- Ladd, D.R. (2008). *Intonational Phonology*. 2<sup>nd</sup> edition. Cambridge: Cambridge University Press.
- Lickley, R., Schepman, A., and Ladd, D.R. (2005). Alignment of “Phrase Accent” Lows in Dutch Falling-Rising Questions: Theoretical and Methodological Implications. *Language and Speech* 48, 157-183.
- Ohala, J. (1983). Cross-language use of pitch: an ethological view. *Phonetica* 40, 1-18.

- Peters, J. and Gussenhoven, C. (ms). *Intonational variation in the Netherlands and beyond. A study of Zeelandic, Hollandic Dutch, West Frisian, Dutch Low Saxon, German Low Saxon, and High German*.
- Smakman, D. (2006). *Standard Dutch in the Netherlands. A sociolinguistic and phonetic description*. PhD thesis. Utrecht: LOT publications.
- 't Hart, J., Collier, R., and Cohen, A. (1990). *A Perceptual Study of Intonation: An Experimental-Phonetic Approach to Speech Melody*. Cambridge: Cambridge University Press.
- van de Ven, M. and Gussenhoven, C. (2011). On the timing of the final rise in Dutch falling-rising intonation contours. *Journal of Phonetics* 39, 225-236.
- van Es, G.A. (1935). Syntactische functies der intonatie in de volkstaal onzer noordelijke provinciën. *Handelingen van het Zestiende Nederlandsche Philologen-Congres*, 39-42.
- van Oostendorp, M. (2002). *Rotterdams*. Den Haag: Sdu Uitgevers.
- Weijnen, A. (1966). *Nederlandse dialectkunde*. Assen: Van Gorcum.
- Wortel, D. (2002). *Leids*. Den Haag: Sdu Uitgevers.



#### Introduction

This dissertation explored how speakers of five local varieties along the Dutch North Sea coast, as well as speakers of Standard Dutch, realize nuclear contours in sentence-medial and sentence-final position. Although it had never been systematically investigated, it has commonly been assumed that Dutch non-tonal dialects have their own characteristic melodies, and that intonation is an important cue to a speaker's linguistic and geographical origins. The present study aimed to find evidence to support this assumption by collecting and analyzing production data from 119 speakers of *Zeelandic Dutch* as spoken in Zuid-Beveland, a *Southern* and *Northern Hollandic* urban vernacular as spoken in Rotterdam and Amsterdam, a *West Frisian* variety spoken in Grou, a *Low Saxon* variety in Winschoten and finally *Standard Dutch* as spoken among the student population of Nijmegen.

Intonation can vary in different ways between languages or language varieties. Ladd (2008:115-116) suggests that cross-linguistic variation in intonation can be broadly described along four dimensions: variation can be categorized as (1) semantic, (2) systemic, (3) phonotactic and (4) realizational. Languages may share the same set of melodies but use them in different situations or to express different meanings (*semantic variation*). It is also possible that a particular melody occurs in one language, but not in another (*systemic variation*). A third category is *phonotactic variation*, which is related to differences in the way languages combine high and low tones to form intonational melodies. Finally, a shared melody may be pronounced slightly differently in two languages, or may be realized differently depending on the context in which it occurs (*realizational variation*).

The primary focus of the current study was on *realizational variation* of intonation. It is known that the pronunciation of structurally identical intonation contours may depend on a variety of contextual factors. Factors considered in this work included the segmental make-up of the word on which a melody is pronounced, the distance of the nuclear accented word to the end of the utterance, the sentence position of the word that carries the accent, and the information structure of the utterance. Apart from contextual variation, earlier literature has also reported more general cross-linguistic differences in the realization of intonation, for example differences in tonal timing and scaling between two varieties of the same language.

In chapters 3 and 4, we described how speakers of the individual dialects responded to the contextual factors *IP-final time pressure* and *focus*. We also investigated whether these responses were different between the dialects. Next, chapter 5 reported what language-specific realizational differences we encountered between the six varieties if contextual factors are disregarded. As such, chapters 3 and 4 were generally concerned with realizational variation *within* the dialects, whereas chapter 5 specifically explored variation *between* the dialects.

The secondary focus of this study was on the remaining three dimensions of variation: semantic, systemic and or phonotactic variation. Based on the data collected for the previous three chapters, chapter 6 reported dialect-specific preferences in contour choice, frequency and distribution. It also presented a detailed description of contour realizations that occurred in one or two dialects only.

Together, the results reported in chapters 3 – 6 aim to answer the question whether dialects have their own intonation ‘signature’. Before answering this question, I will summarize the main results and conclusions per chapter.

### **Summary of chapter 3 – responses to time pressure**

The reading task described in chapter 3 was designed to study how speakers of six varieties of Dutch realized three nuclear contours when they were confronted with time pressure due to (1) the presence of an upcoming IP (intonational phrase) boundary and (2) the complexity of the IP-final nuclear contour. For (1), speakers were asked to read short sentences with a nuclear accent on four IP-final monosyllabic target words (*lof*, *loof*, *lom*, *loom*). We reduced the duration of the sonorant portions in four steps by varying the phonological vowel length of the nucleus (V /ɔ/ vs. VV /o:/) and the sonorant nature of the coda (voiceless /f/ vs. sonorant /m/). The fictitious target words were used as proper names in four statements to elicit H\*L L%, in four yes/no-questions to elicit L\*H H%, L\* H% or H\* H% and in four rhetorical questions to elicit H\*L H%. For (2), we compared the speakers’ realizations of the falls and the fall-rises, which is more complex to produce due to the extra movement. We categorized the participants’ strategies as phonological and phonetic responses, and also wanted to know whether the preferred strategies were dialect- and/or contour-specific.

As for the phonological responses, we expected complex falling-rising contours to be avoided in IP-final position, which none of the varieties structurally did. However, its frequency of use was depressed in the central varieties (in this study: Rotterdam, Amsterdam and Standard Dutch), where the proportion of complex contours was lower for shorter target words.

As for the phonetic responses, all the theoretically possible ways of adjusting the realization of pitch contours were attested in our data. We found truncation,



when a nuclear contour was ended prematurely, most clearly in the case of the nuclear fall and in the final rising movement of the fall-rise for Standard Dutch, Zuid-Beveland, Rotterdam, Amsterdam and Grou. We found compression, when a nuclear contour is produced at higher speed, in the case of the rise in all varieties, and in the nuclear fall and falling part of the fall-rise in Winschoten. Undershoot, when a non-final tone is realized with less extreme  $f_0$ , was applied to the low target of the fall-rise by speakers of Zuid-Beveland and Standard Dutch. Besides changing the shape of the nuclear contour, speakers can create more time to produce them. The varieties of Standard Dutch, Rotterdam and Grou retracted the peak of the fall-rise on words with less voicing, and all varieties increased the duration of segments in fall-rises compared to falls, to facilitate its realization.

Based on these observations, we concluded that both phonetic and phonological adjustments were variety-specific as well as contour-specific. It was not the case, however, that a given dialect had only one type of adjustment for a given contour, nor did we see a stepwise increase in the use of the adjustments going from longer sonorant portions to shorter ones. Instead, speakers resort to a variety of phonetic and phonological adjustment strategies, although some strategies were dominant. For example, speakers of Standard Dutch most clearly responded to IP-final time pressure on fall-rises by truncating the final rise, and additionally employed peak retraction and undershoot of the valley.

The main implication of this result is that it does not suffice to describe languages and contours as either truncating or compressing, which earlier investigations tended to do. Other adjustment strategies or combinations of strategies, phonetic as well as phonological, need to be considered when interpreting responses to time pressure.

An additional, and unexpected, finding of our study was that adjustments to time pressure were not only variety- and contour-specific, but also specific to the phonological context in the case of rises. An analysis of all low rises in Standard Dutch, Rotterdam, Amsterdam and Grou combined showed that vowel class (long/short) and the nature of the coda (fricative/sonorant) affected the alignment of the beginning of the rise. A short vowel in the rime (/ɔ/) resulted in earlier alignment than a long vowel (/o:/). Additionally, the beginning of the rise was earlier in rimes with /m/ compared to rimes with /f/. We currently have no explanation for this finding.

#### **Summary of chapter 4 – effects of information structure**

The objective in this chapter was to examine the effects of focus on the realization of non-final nuclear falls in six varieties. To this end, we designed a reading task with syntactically and lexically identical sentences that were

phonologically ambiguous with respect to the size of the focus constituent (sentence-wide or ‘broad’ vs ‘narrow’) and focus meaning or type (‘informational’ vs ‘corrective’). The declarative sentences favoured a H\*L L% nuclear pitch accent on the non-final accented target words with the metrical pattern sww (*Momberen*, *Memberen*, *Manderen* and *Munderen*). This resulted in four sentences with broad informational focus (BF), four with narrow informational focus (NF) and four with narrow corrective focus (CF). We assumed that CF has a more intensified focus meaning, or a higher ‘information weight’, than NF, while BF has the lowest. The Standard Dutch data set, collected in a pilot experiment, consisted of nine sentences with the target words *Manderen*, *Lunteren* and *Bunderen*. We wanted to know whether the responses to the focus manipulation were variety-specific, and whether responses were different depending of focus *size* or focus *type*.

The results for Standard Dutch were analysed first, and showed that the nuclear pitch accent in narrow focus and corrective focus sentences was realized with shorter and steeper falling slopes, which went down lower and had earlier and slightly lower peaks. In other words, the whole accentual gesture was time-compressed. In addition, the accented syllable was realized with slightly longer onsets and codas in, suggesting that speakers enhanced the pronunciation of the fall as a function of focus *size*, rather than focus *type*.

All the effects that were found in Standard Dutch were also seen in the dialects. Nuclear falls in sentences with higher information weight were realized with shorter, larger and steeper falling slopes, earlier and lower peaks, and longer segmental durations. We also found some additional effects in the rise leading up to the nuclear peak, which had not been attested in Standard Dutch. The elbow preceding the nuclear peak was scaled lower and started earlier, while the rise was larger and longer (or shorter, depending on dialect).

Importantly, only *some* of the responses to the focus manipulation could be observed in *some* of the varieties. Whereas all varieties lowered postfocal *f0*, lower peaks could only be observed in Zuid-Beveland and Amsterdam. The shape (duration, excursion and slope of *f0* movements) was most notably affected in Rotterdam, Amsterdam and Winschoten. Finally, timing effects were observed in Grou and Winschoten, and segmental effects in Winschoten only. Another important outcome was that unlike Standard Dutch, differences in the realization of pitch accents between narrow and corrective focus could be observed just as often as differences between broad and narrow focus, which means our earlier conclusion regarding focus *size* and *type* cannot be extended to the local varieties. Overall, speakers of all varieties tended to use the same strategy to mark higher information weight. Focus overall affected the falling part of the nuclear contour, which was steeper, shorter, went deeper, and sometimes started a little earlier. We interpreted these effects as hyperarticulation, aimed at increasing the

prominence of (parts of) an utterance so as to increase the distinctiveness of different levels of communicative urgency.

Where we found *earlier* peaks and *lowering* of *f0* after the peak, most studies found that larger pitch excursions go hand in hand with *higher* and *later* peaks. Both scenarios are compatible with the hyperarticulation account. If a speaker enhances only the pitch peak, he may do this by raising it and consequently delaying it, resulting in a higher and later peak. Alternatively, if a speaker enhances the entire pitch accent, which contains a nuclear peak, he may do this by a more careful pronunciation of the *f0* fall inside the syllable rime, leading to a steeper fall that may begin earlier and reach lower. What unifies both strategies is that they serve to signal the communicative importance of the pitch accent's focus constituent.

Finally, in light of the objective of this dissertation, we should mention that we haven't been able to identify variety-specific enhancements cues. Rather, we believe that the speaker's goal to hyperarticulate the fall can be attained by using a variety of strategies, using different phonetic cues to different degrees, which together contribute to marking a particular focus structure. Apart from the phonetic cues to mark information structure observed here, the literature also reports cues such as Downstep, deaccentuation, the use of a special pitch accent for focus, or the choice for and number of prenuclear accents. Non-linguistic cues that we may think of are visual clues, body language, eye contact or the shared context between discourse partners.

One unintended contextual clue to information structure in our test materials was the presence of the focus marker 'no' at the start of the corrective focus sentence. This disambiguating morpheme may have had an effect on the necessity for speakers to express the focus structure phonetically. Secondly, our three focus conditions (BF, NF and CF) were presented to the participants in three consecutive blocks. In a short control experiment, this was shown to have an effect on segmental durations in and around the accented word. As a consequence, we were unlikely to find large segmental duration effects. Both aspects should be kept in mind when designing experimental materials for future research.

### **Summary of chapter 5 – dialect-specific realizational differences**

In addition to similar effects to contextual factors, the six varieties also showed language-specific variation in the pronunciation of nuclear contours. This chapter reports dialect-specific pronunciations of IP-medial and IP-final falls, and IP-final fall-rises. It also reported regional variation in the effect of sentence position (final vs. non-final) on the realization of the nuclear accent in falls (H\*L L%).

We used 12 test sentences, all drawn from earlier experiments. Eight were narrow focus declaratives, four with an IP-medial and four with an IP-final falling pitch accent. These were also used for the comparison of non-final and final contours. The remaining four sentences were rhetorical questions with an IP-final falling-rising pitch accent.

As for sentence position, we observed that in all varieties, falls in non-final position were realized with shorter segmental durations, later and higher peaks, shorter *f0* durations, shorter excursions and hence steeper falling slopes. All these effects are reminiscent of those observed in chapter 3 where the effect of time pressure on the realization of IP-final contours was studied. Indeed, there is more time pressure on the realization of a falling contour at the end of an utterance than in sentence-medial position. In other words, the effects observed here for IP-final falls can be attributed to the lack of space to realize the contours. The realization of nuclear contour types was not affected by sentence position uniformly across the dialects. For example, the difference in segmental duration between the non-final and final falling contours was particularly large in Winschoten, and small in Zuid-Beveland. Also, the difference in peak timing between the two falls was much smaller in Winschoten than in Standard Dutch.

Continuing with the effects of dialect on contour realization, we found that *sonorant rime durations* in all three sentence conditions gradually increased from the south-west (Zuid-Beveland) to the north-east (Winschoten). Dialectal differences in *peak timing* were dependent of sentence position. IP-medially, peaks tended to be earliest in Zuid-Beveland and Winschoten, both dialects at the periphery of the Netherlands. IP-finally, Zuid-Beveland still had early peaks, but those in Winschoten were in fact late. We attributed this difference to generous availability of sonorant segments in Winschoten IP-final accented words. Next, the only systematic effect of dialect on *tonal scaling* was observed in the fall-rises, where the valley between the two high tones was realized at much higher *f0* in Zuid-Beveland and Rotterdam than in Amsterdam and Winschoten. The final high tone was scaled highest in Winschoten. Finally, with regard to *contour shape* we observed that the southwestern falls in Zuid-Beveland tended to be small and shallow, in sharp contrast with the large and steep falls in the north-east. South-western fall-rises (Zuid-Beveland) and north-eastern fall-rises (Winschoten) were both realized quite differently from the ‘central’ varieties in Standard Dutch, Rotterdam, Amsterdam and Grou, but each in their own way. In Winschoten, speakers realized the fall-rise with longer, larger and steeper falling movements and shorter and shallower final rises than speakers of the central varieties, leading to a distinct contour shape. In striking contrast with Winschoten, the excursion and slope of both the falling and the final rising movements in Zuid-Beveland were considerably smaller and shallower than in the central varieties. The distinct shape of Zuid-Beveland fall-rises were discussed in more detail in chapter 6, along with final and non-final falling contours of Amsterdam, which stood out

in our comparison with other varieties, due to its speakers producing a combination of normal-peak and late-peak accents. Besides the late timing of the peak, these accents were characterized by particularly long and shallow falling movements.

Significantly, the most extreme realizational variation could be observed in the two geographically most extreme dialects. Zuid-Beveland and Winschoten differed most from each other as well as from the other varieties. They represent two extreme ends of a scale with respect to the phonetic variables reported in this study, with the central varieties of Standard Dutch, Rotterdam, Amsterdam and Grou generally in between.

The main conclusion of this chapter is that the variation in the realization of nuclear contours in the Netherlands seems to follow a geographical cline, with a gradual shift from the south-eastern variety of Zuid-Beveland, via the central varieties of Rotterdam, Amsterdam and Standard Dutch, to the north-eastern varieties of Grou and Winschoten. As such, our data provides additional support for the existence of such a cline, which was first reported in Peters et al. (2014, 2015) for Dutch, Frisian, Low Saxon, and High German intonation.

### **Summary of chapter 6 – non-standard melodies and melodic preferences**

The set of experimental sentences used for this dissertation always consisted of mini-dialogues. Speakers were recorded in pairs, whereby one speaker produced the context sentence and the other the carrier sentence. Even though all mini-dialogues were designed to elicit specific nuclear contours (e.g., declaratives were designed to elicit a falling nuclear pitch accent on a particular accented syllable), speakers sometimes made different choices. Instead of treating those deviant choices as ‘noise’, we decided to look into them in more detail, because they may contribute to a dialect’s intonation ‘signature’.

The first half of this chapter described dialect-specific preferences for nuclear contours in four pragmatic contexts, more specifically in IP-medial declaratives with sentence-wide focus, in IP-final declaratives with narrow focus, in IP-final yes/no-questions with narrow focus and in IP-final rhetorical questions with corrective focus. Looking across pragmatic contexts and across all varieties, we observed eleven different nuclear contours, all of which are known to exist in Standard Dutch. Even though not all eleven contours occurred in all six varieties, this does not imply that some contours are ‘missing’ in a particular dialect. In fact, the missing contours were found in all varieties in Peters and Gussenhoven (ms), who recorded the same set of speakers but whose reading materials included a wider variety of sentence types. On the basis of our results, it is therefore reasonable to conclude that all varieties share a tonal grammar and employ the same tonal contrasts. In other words, we found no evidence of systemic variation.

Interestingly, we found variation in the distribution over sentence types and the frequencies with which all nuclear contours were used. Declaratives were produced mainly with H\*L L% in all varieties. Its use was most frequent in the North-east (West-Frisian and Low Saxon) and least frequent in Rotterdam and particularly Amsterdam, where speakers often used a (delayed) half-completed fall for declaratives. Speakers of Zuid-Beveland used falls (H\*L L%) not only for declaratives, but also in about 80% of yes/no-questions. We concluded that the contour has both a declarative and a strong interrogative meaning in this variety, unlike the other varieties, that all preferred rising intonation for questions. More specifically, about 70-75% of all yes/no-questions in these varieties were produced with simple rises (H\*H%, L\*H H% or L\* H%). Speakers of Standard Dutch, Amsterdam, Grou and Winschoten produced the other 25% with a fall-rise (H\*L H%). The least variation was found in the rhetorical questions, which were realized almost exclusively with H\*L H% in all varieties, although the rise also occurred in Zuid-Beveland, Rotterdam and Amsterdam. Based on these findings, we concluded that although the varieties did not vary systematically, they did vary semantically.

The second half of the chapter was dedicated to three melodies that occurred in specific varieties only: the *interrogative fall* and the *rise-rise* in Zeelandic Dutch as spoken in Zuid-Beveland, and the *(delayed) half-completed fall* from Rotterdam and Amsterdam Dutch.

The dialect of Zuid-Beveland differed from the other varieties in two ways. As mentioned in chapter 3, its speakers often used H\*L L% for yes/no-questions, which other varieties preferred to produce with rises or fall-rises. This *interrogative fall* was shown to be an enhanced version of the *declarative fall*, realized at a higher pitch level, with a larger pitch excursion, a steeper slope and a later peak. Similar cues to enhance the difference between statements and questions have been reported for other languages.

Zuid-Beveland also stood out from the other varieties in its realization of the IP-final fall-rise, which was rather flat in comparison. Rather than a falling-rising melody, it was in fact often shaped as a ‘rising-rising’ melody, having no ‘dip’ between the two high targets. The shorter the IP-final accented word, the larger the proportion of rise-rises was. If a dip was observed, its size varied on a continuum, showing that both rise-rise and fall-rise realizations were phonetic variants of the same phonological contour. We interpreted the flat-shaped fall-rise as a response to IP-final time pressure, where speakers undershot the low target in H\*L H%, sometimes so extremely that it resulted in a rising-rising shape. Based on our data, it is not yet possible to say whether this shape is typical of Zuid-Beveland in general or whether it is only used as a mechanism to deal with time pressure. To find out, future studies ought to collect more data, specifically of IP-medial realizations.

The last dialect-specific contour that was discussed in this chapter were the half-completed fall (H\*L 0%) and the late half-completed fall (L\*HL 0%) used in declaratives sentences with the accented syllable in IP-medial or IP-final position. They were characteristic of Rotterdam and in particular Amsterdam, which used them in around 12% and 44% of all declaratives, respectively. This fall type could be characterized by a slowly descending slope if realized on IP-medial accented words, and by extreme truncation of the falling movement when realized sentence-finally, with more or less delayed peaks. We did not observe a clustering of peak delays around two values and thus no clear distinction between the non-delayed and delayed versions of the half-completed fall, or even between the half-completed and the regular fall (H\*L L%). Yet, the question whether the (late) half-completed falls are phonologically different from the neutral fall cannot be answered convincingly on the basis of our data. One argument in favour of a phonological interpretation is that slow falls combine with normal or late peaks, but rapid falls never combine with late peaks in either of the varieties. On the other hand, the wide range covered by peak timing data points pleads for a phonetic interpretation of the data. Whereas the phonological interpretation would imply that we are dealing with semantic variation, the phonetic interpretation would point in the direction of realizational variation, where the difference lies in the greater degree of variation between Amsterdam and Rotterdam on the one hand, and other varieties on the other. The solution for this unresolved issue may lie in conducting a perception experiment.

The two broad conclusions that were drawn were, first, that both distributional and realizational differences can contribute to a dialect's 'signature', and second that the four ways in which intonation may vary according to Ladd (2008) may not be adequate to describe dialectal differences in intonation. Particularly, Ladd defines semantic differences as variation in the meaning of equatable contours. Dialectal differences in the frequency with which contours are used in the same pragmatic context are captured by this definition only if we allow intonational meaning to be defined by the distributional profile of a contour over some exhaustive set of pragmatic categories. If we were to proceed along that way, we would end up assuming that speakers of different dialects by definition express the same range of meanings in each of these pragmatic conditions. This uncomfortable circularity suggests that determining the communicative effects of intonation contours in different dialects by both native and non-native listeners should be a necessary follow up to research of the sort reported here.

### Final remarks

Let's return to the question raised at the beginning of this dissertation.

*"Do Dutch dialects have their own characteristic intonation, or 'signature'?"*

The short answer is ‘yes, they do’, which is supported by the various examples of variety-specific differences in the realization of intonation contours in different contexts, and by the differences in the distribution of contours over different sentence types. The more differentiated answer is ‘yes, they do, subtly’. The varieties studied here in fact all shared a tonal grammar and set of tonal contrasts. Their responses to contextual manipulations were often in the same direction, making use of the same set of possible cues, but with subtle differences in which cues were a particular dialect preferred. More obvious realizational differences could be observed if dialects were directly compared, disregarding context. We found evidence of a geographical cline of phonetic variation from the southwest to the northeast, particularly regarding pitch excursion sizes and segmental durations. Finally, we observed clear differences in the melodic preferences that varieties had in different pragmatic contexts, and discovered that some nuclear contours in some varieties were realized quite radically different from the other varieties’ pronunciations. These non-standard realizations did not fit the geographical cline, and, along with the distributional preferences, perhaps form the largest contribution to a dialect’s signature.

Of course, we have probably unravelled only a small part of the intonational signatures, perhaps their initials, leaving their full names to be discovered. Future research on variation in dialect intonation should therefore focus on investigating variation (particularly of the distributional sort) in other pragmatic contexts than only IP-final and IP-medial falls, and IP-final rises and fall-rises. For example, we have not yet looked at the realization of prenuclear accents. Moreover, the investigations should target more varieties, to further look into the claim of a geographical cline. It would be interesting to see whether patterns are sensitive to political boundaries.

It will also be worth investigating whether some of the more salient phonetic differences, such as the Zuid-Beveland ‘rise-rise’ and the Amsterdam late-peak accents, can be recognized as a regional characteristic by indigenous and endogenous speakers. Perception experiments will also be useful to find out whether the late peak and normal peak pitch accents are indeed phonetic variants of the same phonological category, as we suggested, or whether they represent two different pitch accents.

One specific result which deserves further exploration is the fact that in our data, vowel class and the nature of the coda affected the alignment of the beginning of the rise. We could not explain this finding, but recommend that future research focuses on these segmental effects.

In general, we would advise future researchers to record speakers from different social backgrounds. Possibly, our participants’ speech (generally well-educated, middle class) may have been rather influenced by Standard Dutch.



APPENDIX

Reader’s guide:

- A context sentence
- B carrier sentence, target word in **bold** (only for Dutch)
- [fd] left and right edge of focus domain (only for Dutch and only in set 4)
- i. Dutch (Zuid-Beveland, Rotterdam, Amsterdam)
- ii. West Frisian (Grou)
- iii. Low Saxon (Winschoten)
- iv. English gloss
- v. *English translation*
- vi. Standard Dutch pilot

Where materials differ between the varieties, separate English glosses and translations are provided for each, indicated by slashes (/).

The shaded columns at the right of the table indicate in which chapters (3, 4, 5 or 6) the experimental sentences are used.

Set 1: Statements with IP-final pitch accent (4)

Sentence 1		3		5	6
A	i. Met wie gaat Roel naar 't concert?				
	ii. Mei wa sil Roel nei 't konsert?				
	iii. Mit wel gaait Roel noar 't concert?				
	iv. with who goes Roel to the concert				
	v. <i>With whom is Roel going to the concert?</i>				
B	i. Hij gaat met Marjel de <b>Lof</b> .				
	ii. Hy sil mei Marjel de Lof.				
	iii. Hai gaait mit Marjel de Lof.				
	iv. he goes with Marjel de Lof				
	v. <i>He's going with Marjel de Lof.</i>				

Sentence 2		3		5	6
A	i. Van wie is dat dikke boek?				
	ii. Fan wa is dat tsjûke boek?				
	iii. Van wel is dat dikke bouk?				
	iv. of who is that big book				
	v. <i>Whose big book is that?</i>				
B	i. Dat is van Professor <b>Loof</b> .				
	ii. Dat is fan Professor Loof.				
	iii. Dat is van Professor Loof.				
	iv. that is of professor Loof				
	v. <i>It belongs to professor Loof.</i>				
Sentence 3		3		5	6
A	i. Met wie gingen de kinderen naar de dierentuin?				
	ii. Mei wa gienen de bern nei de dieretún?				
	iii. Mit wel gingen de kiender noar de daierntoene?				
	iv. with who went the children to the zoo				
	v. <i>Who did the children go to the zoo with?</i>				
B	i. Ze gingen met meester <b>Lom</b> .				
	ii. Se gienen mei meester Lom.				
	iii. Zai gingen mit meester Lom.				
	iv. they went with mister Lom				
	v. <i>They went with mister Lom.</i>				
Sentence 4		3		5	6
A	i. Met wie gaat je baas morgen trouwen?				
	ii. Mei wa sil dyn baas moarn trouwe?				
	iii. Mit wel gaat dien boas mörgen traauwen?				
	iv. with who goes your boss tomorrow marry				
	v. <i>Who will your boss marry tomorrow?</i>				
B	i. Hij trouwt met mevrouw de <b>Loom</b> .				
	ii. Hy trout mei mefrou de Loom.				
	iii. Hai traauwt mit vraauw de Loom				
	iv. he marries with mrs de Loom				
	v. <i>He'll marry Mrs. de Loom.</i>				

vi.	Standard Dutch (pilot), n=4	3		5	6
1	A Met wie gaat Roel naar 't concert? B Hij gaat met meneer 't Lof.				
2	A Van wie is dat dikke boek? B Dat is van Professor Loof.				
3	A Met wie gingen de kinderen naar de dierentuin? B Ze gingen met meester Lom.				
4	A Met wie gaat je baas morgen trouwen? B Hij trouwt met mevrouw De Loom.				

### Set 2: Yes/no-questions with IP-final pitch accent (4)

Sentence 1	3			6
A i. Trek je jas aan; we vertrekken zo naar Eindhoven. ii. Doch de jas oan; wy ferreizgje sa nei Eindhoven. iii. Trek dien jak aan; wie vertrekken zo noar Eindhoven.  iv. pull your coat on; we leave shortly to Eindhoven v. <i>Put on your coat; we're about to go to Eindhoven.</i>				
B i. Gaan we daar naar tante <b>Lof</b> ? Gezellig! ii. Sille wy dêr nei tante Lof? Gesellich! iii. Goan wie doar noar tante Lof? Behoagelk!  iv. go we there to aunt Lof? nice v. <i>Are we going to aunt Lof? How nice!</i>				
Sentence 2	3			6
A i. We gaan zo onze caravan ophalen voor de vakantie. ii. Wy sille sa ús caravan ophelje foar de fakânsje. iii. Wie goan zo onze caravan ophoalen veur de vekansie.  iv. we go shortly our caravan pick up for the holiday v. <i>We're about to pick up our caravan for the holidays.</i>				
B i. Staat-ie bij familie <b>Loof</b> ? Daar stallen wij 'm ook altijd. ii. Stiet er by famylje Loof? Dêr stalje wy 'm ek altyd. iii. Staait e bie fermilie Loof? Doar stalen wie hom ook aaltied.  iv. stands it at family Loof? there put up we it also always v. <i>Is it at the Loof family? We always park it there as well.</i>				

Sentence 3		3		6
A	i. Ik zag net je broer Koen met je buurvrouw langslopen. ii. Ik seach niis dyn broer Koen mei dyn buorfrou delrinnen. iii. Ik zag net dien bruier Koen mit dien noaberwief langs lopen.  iv. i saw just your brother Koen with your neighbor walk by v. <i>I just saw your brother Koen walk by, with your neighbor.</i>			
	B i. Liep-ie naast mevrouw de Lom? Wat raar, die kennen elkaar toch niet? ii. Rûn er neist mefrou de Lom? Wat raar, dy kenne elkoar dochs net? iii. Laip e noast vraauw de Lom? Wat roar, dij kennen mekoar toch nait?  iv. walked he next mrs de Lom? what strange, they know each other actually not? v. <i>Did he walk next to Mrs. de Lom? How strange, they don't know each other, do they?</i>			
Sentence 4		3		6
A	i. Ik moet straks naar de baas komen, omdat ik weer te laat was vanmorgen. ii. Ik moat strak by de baas komme, omdat ik wer te let wie fan 'e moarn. iii. Ik mout straks noar de boas komen, omdat ik vanmörgeen weer te loat was.  iv. i must later to the boss come, because i again too late was this morning v. <i>The boss wants to see me later today, because I was late again this morning.</i>			
	B i. Moet je dan naar dokter Loom? Oei, maak je borst dan maar nat! ii. Moatsto wer nei dokter Loom? Oei, meitsje dy dan mar klear! iii. Most doe noar dokter Loom? Oei, moak dien borst din moar nat!  iv. must you then to doctor Loom? ouch, make your chest then just wet! v. <i>Are you seeing doctor Loom? Oh no, you'd better brace yourself!</i>			
vi.	Standard Dutch (pilot), n=4	3		6
1	A 'Trek je jas aan; we vertrekken zo naar Eindhoven. B Gaan we daar naar tante Lof? Gezellig!			
2	A We gaan zo onze caravan ophalen voor de vakantie. B Staat-ie bij familie Loof? Daar stallen wij 'm ook altijd.			
3	A Volgens mij zag ik net je oude basisschool leraar. B Zag je meneer Van der Lom? Wat leuk, die heb ik al jaren niet gezien.			
4	A Ik moet straks naar de baas komen, omdat ik weer te laat was vanmorgen. B Moet je naar meneer De Loom? Oei, maak je borst dan maar nat!			

### Set 3: Rhetorical questions with IP-final pitch accent (4)

Sentence 1		3		5	6
A	i. Ga je mee naar Bakkerij 't Stoepje?				
	ii. Giest do mei nei Bakkerij 't Stoepje?				
	iii. Gaaist doe mit noar Bakkerij 't Stoepje?				
	iv. go you along to bakery 't Stoepje?				
	v. <i>Do you want to come along to Bakery 't Stoepje?</i>				
B	i. We gaan toch naar Bakker <b>Lof</b> ?				
	ii. Wy gean dochs nei Bakker Lof?				
	iii. Wie goan toch noar Bakker Lof?				
	iv. we go actually to baker Lof?				
	v. <i>But aren't we going to Baker Lof?</i>				
Sentence 2		3		5	6
A	i. Meester Boelens gaat mee op schoolreis.				
	ii. Master Boelens sil mei op skoalreis.				
	iii. Meester Boelens gaait mit op schoulraaise.				
	iv. mister Boelens goes along on schooltrip				
	v. <i>Mister Boelens is coming along on the school trip.</i>				
B	i. Je ging toch met meester <b>Loof</b> ?				
	ii. Do giest dochs mei master Loof?				
	iii. Doe gingst toch mit meester Loof?				
	iv. you went actually with mister Loof?				
	v. <i>But weren't you going with Mister Loof?</i>				
Sentence 3		3		5	6
A	i. Pepijn de Heer komt straks ook naar 't feest.				
	ii. Pepijn de Heer komt strak ek nei 't feest.				
	iii. Pepijn de Heer komt straks ook noar 't feest.				
	iv. Pepijn de Heer comes later also to the party				
	v. <i>Pepijn de Heer is also coming to the party later.</i>				
B	i. Hij heet toch Pepijn de <b>Lom</b> ?				
	ii. Hy hjit dochs Pepijn de Lom?				
	iii. Hai hait toch Pepijn de Lom?				
	iv. he calls actually Pepijn de Lom?				
	v. <i>But isn't he called Pepijn de Lom?</i>				

Sentence 4		3		5	6
A	i. Dit antieke horloge is nog van opa Thijssen geweest.				
	ii. Dit antike horloazje hat noch fan pake Thijssen west.				
	iii. Dit antieke hallozie is nog van opa Thijssen west.				
	iv. this antique wristwatch is still from grandfather Thijssen been				
	v. <i>This antique wristwatch used to belong to grandfather Thijssen.</i>				
B	i. Het was toch van opa <b>Loom</b> ?				
	ii. It wie dochs fan pake Loom?				
	iii. Het was toch van opa Loom?				
	iv. it was actually from grandfather Loom?				
	v. <i>But didn't it belong to grandfather Loom?</i>				
vi. Standard Dutch (pilot), n=4		3		5	6
1	A Ga je mee naar Café de Engel?				
	B Je wilde toch naar 't Lo?				
2	A Meester Boelens gaat mee op schoolreis.				
	B Je ging toch met meester Loof?				
3	A Mies de Heer komt straks ook naar 't feest.				
	B Ze heette toch Mies de Lom?				
4	A Dit antieke horloge is nog van opa Thijssen geweest.				
	B Het was toch van opa Loom?				

**Set 4: Statements with IP-medial pitch accent (12)**

Wide informational focus – BF					
Sentence 1			4		6
A	i. Zijn er al vakantieplannen?				
	ii. Binne der al fakânsjeplannen?				
	iii. Binnen der al vakantieplannen?				
	iv. are there already holidayplans				
	v. <i>Are there any plans for the holidays?</i>				
B	i. [We zouden naar <b>Munderen</b> willen liften.]				
	ii. Wy soenen nei Munderen liftsje wolle.				
	iii. Wie zollen noar Munderen liften willen.				
	iv. we would to munderen want hitchhike / hitchhike want				
	v. <i>We'd like to hitch a ride to Munderen.</i>				

Sentence 2		4	6
A	i. Zijn er voor vanavond al plannen gemaakt? ii. Binne der foar jûn al plannen makke? iii. Binnen der veur vanoavond al plannen moakt?  iv. are there for tonight already plans made v. <i>Are there any plans for tonight?</i>		
B	i. Ja, [we zouden bij <b>Memberen</b> willen eten.] ii. Ja, wy soenen by Memberen ite wolle. iii. Joa, wie zollen bie Memberen eten willen.  iv. yes we would at memberen want eat / eat want v. <i>Yes, we'd like to eat at Memberen's.</i>		
Sentence 3		4	6
A	i. Wat zijn de plannen voor morgen? ii. Wat binne de plannen foar moarn? iii. Wat binnen de plannen veur mörgen?  iv. what are the plans for tomorrow v. <i>What are the plans for tomorrow?</i>		
B	i. [Ik zou wel naar <b>Momberen</b> willen fietsen.] ii. Ik soe wol nei Momberen fytse wolle. iii. Ik zol wel noar Momberen lopen willen.  iv. i would quite to momberen want cycle / cycle want / walk want v. <i>I'd quite like to cycle / walk to Momberen.</i>		
Sentence 4		4	6
A	i. Wat is er? ii. Wat is der? iii. Wat is ter?  iv. what is there v. <i>What's up?</i>		
B	i. [Ik had je naar <b>Manderen</b> willen rijden.] ii. Ik hie dy nei Manderen ride wollen. iii. Ik zol die oet Manderen hoalen willen.  iv. i had you to manderen want drive / drive want // i will you out manderen get want v. <i>I'd wanted to drive you to Manderen // I wanted to pick you up from Manderen</i>		

Narrow informational focus – NF					
Sentence 5			4	5	6
A	i. Waar zouden jullie willen blijven? ii. Wêr soenen jimme bliuwe wolle? iii. Woar zollen ie langes willen?  iv. where would you want stay / stay want // where would you by want  v. <i>Where would you like to stay? // Where would you like to pass by?</i>				
B	i. We zouden bij <b>Munderen</b> willen blijven. ii. Wy soenen by Munderen bliuwe wolle. iii. Wie zollen bie Munderen langes willen.  iv. we would at munderen want stay / stay want // we would at munderen by want v. <i>We'd like to stay at Munderen // We'd like to pass by Munderen.</i>				
Sentence 6			4	5	6
A	i. Waar zouden de Janssens heen willen lopen? ii. Wêr soenen de Janssens hinne rinne wolle? iii. Woar zollen de Janssens hinlopen willen?  iv. where would the johnsons to want walk / walk want v. <i>Whereto would the Johnsons like to walk?</i>				
B	i. Ze zouden naar <b>Memberen</b> willen lopen. ii. Se soenen nei Memberen rinne wolle. iii. Zai zollen noar Memberen lopen willen.  iv. they would to memberen want walk / walk want v. <i>They'd like to walk to Memberen.</i>				
Sentence 7			4	5	6
A	i. Waar zou Karel je heen willen brengen? ii. Wêr soe Karel dy hinne bringe wolle? iii. Woar zol Karel die hinbringen willen?  iv. where would karel you to want bring / bring want v. <i>Where did Karel want take you to?</i>				
B	i. Hij zou me naar <b>Momberen</b> willen brengen. ii. Hy soe my nei Momberen bringe wolle. iii. Hai zol mie noar Momberen bringen willen.  iv. he would me to momberen want bring / bring want v. <i>He wanted to take me to Momberen.</i>				



Sentence 8		4	5	6
A	i. Waar zouden je oom en tante willen wonen?			
	ii. Wêr soenen jîm omke en tante wenje wolle?			
	iii. Woar zollen dien oom en tante wonen willen?			
	iv. where would your uncle and aunt want live / live want			
	v. <i>Where would your uncle and aunt like to live?</i>			
B	i. Ze zouden bij <b>Manderen</b> willen wonen.			
	ii. Se soenen by Manderen wenje wolle.			
	iii. Ze zollen bie Manderen wonen willen.			
	iv. they would at manderen want live / live want			
	v. <i>They'd like to live near Manderen.</i>			
Narrow corrective focus – CF				
Sentence 9		4		6
A	i. Had je me uit Beverwijk willen halen?			
	ii. Hiesto my út Beverwijk helje wollen?			
	iii. Zolst doe mie oet Beverwijk hoalen willen?			
	iv. had you me out beverwijk want get / get want			
	v. <i>Did you want to pick me up from Beverwijk?</i>			
B	i. Nee, ik had je uit [ <b>Munderen</b> ] willen halen.			
	ii. Nee, ik hie dy út Munderen helje wollen.			
	iii. Nee, ik zol die oet Munderen hoalen willen.			
	iv. no I had you out munderen want get / get want			
	v. <i>No, I wanted to pick you up from Munderen.</i>			
Sentence 10		4		6
A	i. Zouden de piloten bij Eindhoven willen landen?			
	ii. Soenen de piloaten by Eindhoven lânje wolle?			
	iii. Zollen de piloten bie Eindhoven landen willen?			
	iv. would the pilots at eindhoven want land / land want			
	v. <i>Would the pilots want to land at Eindhoven?</i>			
B	i. Nee, ze zouden bij [ <b>Memberen</b> ] willen landen.			
	ii. Nee, se soenen by Memberen lânje wolle.			
	iii. Nee, ze zollen bie Memberen landen willen.			
	iv. no they would at memberen want land / land want			
	v. <i>No, they want to land at Memberen.</i>			

Sentence 11		4	6
A	i. Zouden jullie naar Montfort willen rennen? ii. Soenen jimme nei Montfort rinne wolle? iii. Zollen joe noar Montfort lopen willen?  iv. would you to montfort want run / run want / walk want v. <i>Would you like to run / walk to Montfort?</i>		
	B i. Nee, we zouden naar <b>[Manderen]</b> willen rennen. ii. Nee, wy soenen nei Manderen rinne wolle. iii. Nee, wie zollen noar Manderen lopen willen.  iv. no we would to manderen want run / run want / walk want v. <i>No, We'd like to run / walk to Manderen.</i>		
Sentence 12		4	6
A	i. Had je moeder je naar Zaltbommel willen sturen? ii. Hie jimme mem dy nei Zaltbommel stjoere wollen? iii. Zol dien moeke in Zaltbommel wonen willen?  iv. had your mother you to zaltbommel want send / send want // would your mother in zaltbommel live want v. <i>Did your mother want to send you to Zaltbommel? // Would your mother like to live in Zaltbommel?</i>		
	B i. Nee, ze had me naar <b>[Momberen]</b> willen sturen. ii. Nee, se hie my nei Momberen stjoere wollen. iii. Nee, ze zol in Momberen wonen willen.  iv. no she had me to momberen want send / send want // no she would in momberen live will v. <i>No, she wanted to send me to Momberen // No, she'd like to live in Momberen.</i>		
vi. Standard Dutch (pilot)			
Broad focus		4	6
1	A Wat wil de familie Mols morgen doen? B Ze zouden naar Bunderen willen fietsen.		
	2 A En, zijn jullie van 't weekend nog weggeweest? B Ja, we waren in Lunteren gaan logeren.		
3	A Wat is er met jullie? B We willen in Manderen blijven wonen.		

<b>Narrow focus</b>		<b>4</b>	<b>5</b>	<b>6</b>
4	A Waar zou de familie Mols heen willen fietsen? B Ze zouden naar Bunderen willen fietsen.			
5	A Waar waren jullie gaan logeren? B We waren in Lunteren gaan logeren.			
6	A Waar willen jullie blijven wonen? B We willen in Manderen blijven wonen.			
<b>Corrective focus</b>		<b>4</b>		<b>6</b>
7	A Zou de familie Mols naar Zaltbommel willen fietsen? B Nee, ze zouden naar Bunderen willen fietsen.			
8	A Dus jullie waren in Zutphen gaan logeren? B Nee, we waren in Lunteren gaan logeren.			
9	A Willen jullie in Montfort blijven wonen? B Nee, we willen in Manderen blijven wonen.			



## SAMENVATTING IN HET NEDERLANDS

---

Sprekers uit Limburg, waar de meeste dialecten woordtoon kennen, zijn gemakkelijk te herkennen aan hun intonatie. Het wordt algemeen aangenomen dat ook andere varianten van het Nederlands hun eigen, karakteristieke melodieën hebben, en dat je op basis van intonatie kunt weten waar sprekers vandaan komen. Deze aanname blijkt uit het feit dat dialecten ‘zangerig’ worden genoemd, of juist ‘monotoon’.

In deze dissertatie is onderzocht of er inderdaad verschil is in de uitspraak van zinsmelodieën. Daarvoor hebben ca. 120 sprekers van het Zeeuws, het Rotterdams, het Amsterdams, het West-Fries, het Nedersaksisch en het Standaardnederlands meegedaan aan een taalproductie-experiment.

Intonatie bestaat uit hoge en lage tonen. Door deze tonen op bepaalde woorden uit te spreken, kun je als spreker duidelijk maken welke informatie belangrijk is. Zulke ‘toonhoogteaccenten’ vormen samen met de tonen aan het begin en eind van zinnen een zinsmelodie, die betekenis draagt (bv. vraagintonatie). Talen kunnen op een aantal manieren verschillen in intonatie laten zien. Mogelijk gebruiken sprekers van twee talen dezelfde melodie, maar heeft deze in de ene taal een andere betekenis dan in de andere. Het komt ook voor dat een melodie wel in de ene, maar niet in de andere taal bestaat, of dat talen verschillen in de combinaties van hoge en lage tonen die mogelijk zijn. Tot slot is het mogelijk dat twee talen over dezelfde set melodieën beschikken, maar dat deze in de ene taal net iets anders worden uitgesproken (iets hoger of iets lager of met iets langer aangehouden tonen) dan in de andere. Dit onderzoek richt zich met name op dit laatste type variatie, ofwel uitspraakverschillen. De vraag wordt beantwoord of er inderdaad zoiets bestaat als een typische dialectintonatie.

Nieuwe of belangrijke informatie in een zin wordt meestal uitgedrukt met behulp van een toonhoogteaccent, dat valt op de beklemtoonde lettergreep, bijvoorbeeld in *Felix wil graag nieuwe LEGO kopen*. Als juist het woord ‘nieuw’ belangrijk is, valt daar ook het toonhoogteaccent op: *Felix wil graag NIEUWE lego kopen*. Soms is onduidelijk welke informatie in de zin nieuw is, zonder de context te weten. Als antwoord op de vraag *Wat wil Felix kopen?* is ‘nieuwe lego’ de nieuwe informatie, maar als antwoord op de vraag *Wil Felix nieuwe boeken kopen?* is alleen het woord ‘lego’ de nieuwe informatie. Toch valt hier in beide antwoorden het toonhoogteaccent op het woord lego. We weten dat in sommige talen het toonhoogteaccent iets anders wordt uitgesproken in zulke ambigue gevallen en hebben dat ook voor de dialecten in Nederland onderzocht.

Uit de resultaten bleek dat dalende toonhoogteaccenten op informatie die als ‘belangijker’ of ‘urgenter’ kan worden gekenmerkt, werden uitgesproken op wat langere lettergrepen, met iets lagere en vroegere pieken, met kortere en steilere dalingen die een lager punt bereikten. De meeste van deze effecten zijn ook in andere talen aangetroffen. Geen van de dialecten laat alle mogelijke aanpassingen zien, noch is er in bepaalde dialecten een sterke voorkeur voor één bepaalde aanpak. In het algemeen zijn de aanpassingen erop gericht om aan sommige gedeeltes van de zin meer nadruk te geven dan aan andere en zo verschillende niveaus van ‘urgentie’ te kunnen uitdrukken.

Het is bekend dat bijvoorbeeld dalende melodieën anders worden uitgesproken afhankelijk van de afstand tot het zinseinde. Zo is een toonhoogtedaling in *‘Oskar is naar de biosCOOP geweest’* anders dan die in *‘Oskar was in de biosCOOP’*, waar deze op één korte lettergreep uitgesproken moet worden. We hebben gekeken naar het effect van zulke tijdsdruk op de uitspraak van drie soorten melodieën (dalingen, stijgingen en daal-stijgingen) in de verschillende dialecten. Meer complexe melodieën zoals daal-stijgingen (als in *‘Tygo wil toch OOK mee naar buiten?’*) worden in sommige talen vermeden op een enkele lettergreep aan het eind van de zin. Dat was niet het geval in onze data: daal-stijgingen werden in alle dialecten gewoon uitgesproken, al werden ze op hele korte woorden in Rotterdam, Amsterdam en in het Standaardnederlands soms vervangen door een stijgende melodie.

Ook al gebruikten de meeste sprekers alle melodieën ook op de laatste lettergrepen van een zin, ze pasten die wel aan onder de ontstane tijdsdruk. Bekende strategieën in de literatuur zijn het inkorten van de melodie (truncatie) of het sneller uitspreken ervan (compressie). Sprekers kunnen ook eerder beginnen of bijvoorbeeld het verschil tussen de lage en de hoge toon (de excursie) kleiner maken. Al deze strategieën zijn we tegengekomen in het onderzoek, maar niet alle strategieën bij alle dialecten. De conclusie is dat de traditionele indeling in talen die ‘trunceren’ en talen die ‘comprimeren’ niet volstaat en dat de strategieën die sprekers inzetten bij tijdsdruk complexer zijn. Sprekers van hetzelfde dialect gebruiken een combinatie van meerdere strategieën en wisselen ertussen, al toonden ze wel voorkeuren voor een bepaalde aanpak. Zo trunceren sprekers van het Standaard Nederlands, het Rotterdams, Amsterdams en West-Fries de laatste stijging in daal-stijgingen. Sprekers uit Winschoten comprimeren het eerste gedeelte van de daal-stijging, terwijl sprekers van het Zeeuws vooral de lage toon in het midden minder laag uitspreken. Deze toon wordt in veel gevallen zelfs zo hoog, dat er van een daal-stijging een hoorbare ‘stijg-stijging’ wordt gemaakt. Daarmee onderscheidt het Zeeuws zich in deze context duidelijk van de andere dialecten.

In onze data hebben we elf verschillende intonatiecontouren aangetroffen. Het is aannemelijk dat alle onderzochte dialecten dezelfde tonale grammatica hebben.

Wel verschillen de dialecten in de keuze voor bepaalde melodieën of de frequentie waarmee ze deze melodieën gebruiken.

Waar de overige dialecten bijvoorbeeld kiezen voor een stijgende melodie of een daal-stijging om ja/nee-vragen mee uit te spreken, gebruikt het Zeeuws een daling. Grammaticaal gezien is dat dezelfde intonatie als ze voor stellende zinnen gebruiken. Wel wordt deze ‘vragende’ daling anders uitgesproken: de toonhoogte is hoger, de daling is groter, steiler en de piek van het toonhoogteaccent is later. Deze kenmerken om vragen van stellingen te onderscheiden, komen ook in andere talen voor.

Ook in het Rotterdams en Amsterdams komen we verschillen met de andere dialecten tegen. Deze varianten van het Nederlands gebruiken voor stellende zinnen vaak een daling waarvan de piek pas in de lettergreep na de klemtoon wordt uitgesproken en die vervolgens heel langzaam daalt naar het zinseinde. Bij de andere dialecten wordt de piek van het toonhoogteaccent uitgesproken op de beklemtoonde lettergreep, waarna de daling snel wordt ingezet. De typisch Amsterdamse en Rotterdamse uitspraak komt wel voor in reclames. De ‘late-piek’-melodie kent buiten Amsterdam en Rotterdam dus een andere betekenis.

Kijken we tot slot naar meer algemene uitspraakverschillen tussen de wijze waarop melodieën in de verschillende dialecten worden uitgesproken, dan lijkt geografie daarbij een rol te spelen. Over het algemeen vormen het zuidoosten (Zeeuws) en het noordoosten (Nedersaksisch) niet alleen geografisch de twee uitersten van het spectrum, maar ook als we kijken naar uitspraak van intonatie. In Zeeland zijn de lettergrepen waarop de melodie wordt uitgesproken het kortst, en in Winschoten het langst. Het Zeeuws kent daarnaast vroege pieken, met minder extreme verschillen in toonhoogte, en minder steile melodieën. In het Nedersaksisch zijn de pieken meestal later, de toonhoogteverschillen groter, en de melodieën steiler. De overige, meer centrale dialecten, van Rotterdam tot het West-Fries, vallen wat betreft uitspraak meestal tussen deze extremen. Zo worden de duren van de lettergreep bijvoorbeeld geleidelijk aan langer van het zuidoosten naar het noordoosten. Voor andere kenmerken, zoals de neiging om de daling later te beginnen, blijken de centrale dialecten zich van de dialecten ten noorden en ten zuiden te onderscheiden.

Op basis van dit onderzoek kan geconcludeerd worden dat dialecten in Nederland inderdaad typische intonatiekenmerken hebben, maar dat de verschillen over het algemeen klein zijn. De dialecten maken gebruik van dezelfde tonale grammatica maar maken andere keuzes in de toepassing daarvan in verschillende contexten. Ook zijn er kleine uitspraakverschillen. De dialecten die aan de grenzen van Nederland gesproken worden, verschillen daarbij het sterkst van elkaar. Het zou interessant zijn om te onderzoeken of deze geografische tendens doorzet als er verder wordt gekeken naar verwante dialecten die in België en Duitsland worden gesproken.





## CURRICULUM VITAE

---

Judith Hanssen was born in the Netherlands on 17 December 1979. After obtaining her Gymnasium diploma from College Sittard in 1997 and spending six months in Aberystwyth, Wales, she started studying English language and literature in Nijmegen in 1999. In her third year, she received a Harting Scholarship. This allowed her to study Linguistics and teach Dutch at Reading University, UK, in 2001-2002. In 2005, she graduated *cum laude* with an MA thesis on the dialect of Sittard. From January 2006 to December 2010, she worked as a PhD student at the Centre for Language Studies at the Radboud University Nijmegen, which resulted in this dissertation. Currently, Judith holds a position as a lecturer at Avans University of Applied Sciences in 's Hertogenbosch, the Netherlands.